Taggable Pyrrolysine Analogs

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation
with research distinction in Chemistry in the undergraduate colleges
of The Ohio State University

by
Kathleen E. McCormick

The Ohio State University
June 2009

Project Advisor: Professor Michael K. Chan, Departments of Chemistry and Biochemistry

Acknowledgements

Assistance with biochemical studies from Xin Li and Marianne M. Lee, and financial support from The Ohio State University are gratefully acknowledged.

Abstract

N^6-(1,3-dioxolancarbonyl)lysine was synthesized and incorporated into full-length mCherry and α-synuclein proteins by reading through a UAG codon. Mass spectrometry analysis confirmed that the analog incorporates into α-synuclein in the acetal form. Two additional pyrrolysine analogs, N^6-(2-oxo-2*H*-chromene-3-oyl)lysine trifluoroacetate and N^6-(aminoxyacetoyl)lysine ditrifluoroacetate, were synthesized but neither compound reads through the UAG codon.

Table of Contents

1. Introduction

Pyrrolysine (**1**, Figure 1), the 22nd genetically-encoded amino acid, was discovered by the laboratory of Professor Michael K. Chan in collaboration with the laboratory of Professor Joseph Krzycki in 2002.[1,2] Pyrrolysine has since been incorporated into recombinant proteins in *E. coli* by reading through the UAG codon.[3] In addition, various pyrrolysine analogs (**2**-**6**, Figure 1) have been successfully synthesized, some of which (**2**, **4**-**6**) have been incorporated into recombinant proteins.[4,5]

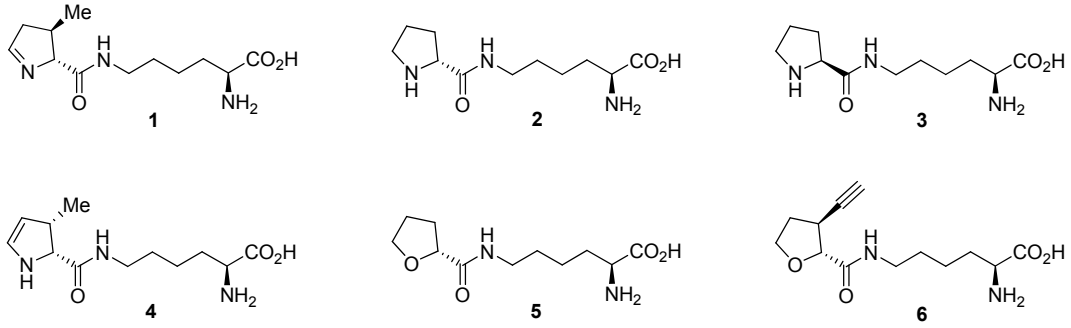
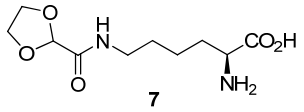


***Figure 1**: Pyrrolysine (**1**) and its synthetic analogs (**2**-**6**).*

In order to attempt to introduce an aldehyde functionality into a protein system, a new target pyrrolysine analog, N^6-(1,3-dioxolancarbonyl)lysine (**7**), was chosen.



This analog was chosen because of its steric and electronic similarity to pyrrolysine, chemical stability, and lack of a chiral center at the N^6 substituent of lysine. In addition, this target analog contains an acetal which potentially could be converted to an aldehyde under acidic conditions. Once the aldehyde is unmasked, a hydrazine dye could be used to tag the protein via hydrazone formation. Areas of specific proteins, such as α-synuclein, which appear to have salt bridges between neighboring lysine and glutamate residues could then be studied. In order to

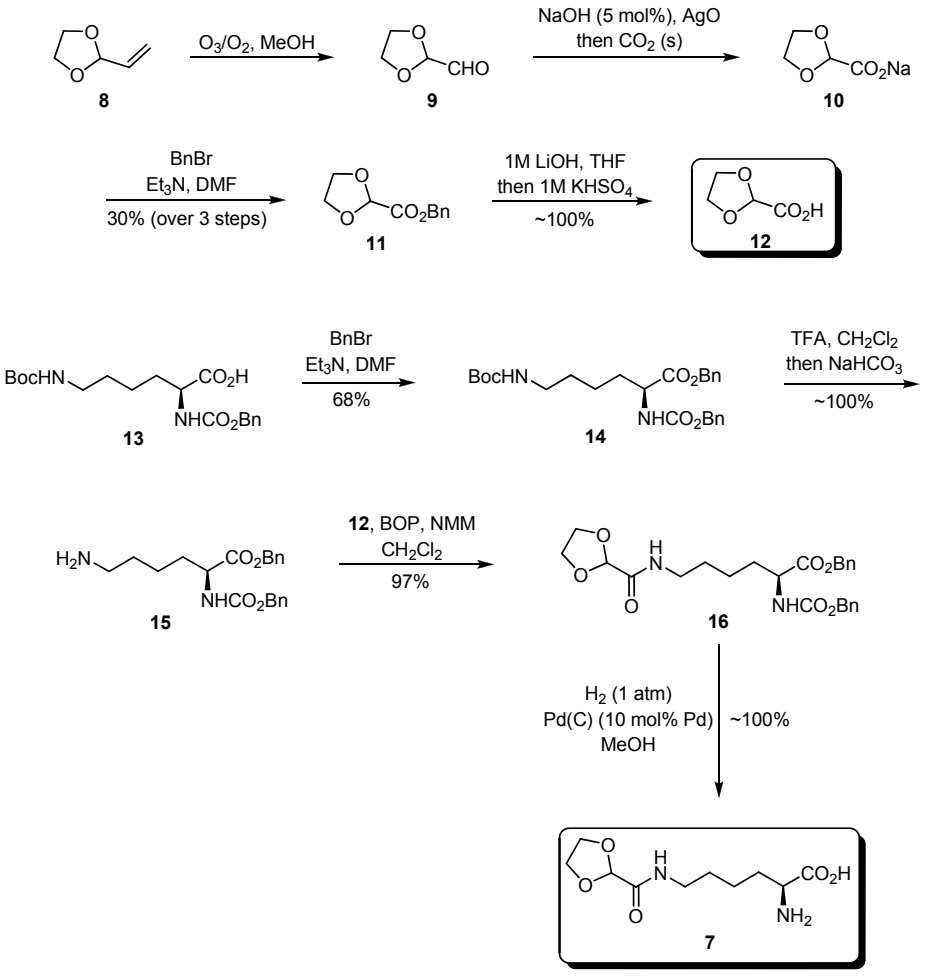
[1]G. Srinivasan, C. M. James, J. A. Kryzycki, *Science* **2002**, *296*, 1459–1462.
[2]B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462–1466.
[3]S. K. Blight, R. C. Larue, A. Mahapatra, D. G. Longstaff, E. Chang, G. Zhao, P. T. Kang, K. B. Green-Church, M. K. Chan, J. A. Krzycki, *Nature* **2004**, *431*, 333–335.
[4]C. R. Polycarpo, S. Herring, A. Berube, J. L. Wood, D. Soll, A. Ambrogelly, *FEBS* **2006**, *580*, 6695–6700.
[5]T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem. Int. Ed.* **2009**, *48*, 1633–1635.

determine if a salt bridge occurs in these areas, specific glutamate residues could be replaced with the synthetic pyrrolysine analog **7**.

2. Results, Discussion, and Conclusions: Organic Chemistry Studies

2.1. Taggable Pyrrolysine Analog (7)

Pyrrolysine analog **7** was synthesized with an overall yield of 20% as depicted in Scheme 1. The first three steps were performed without isolating or purifying the intermediates (**8**-**10**) to obtain ester **11**, which was then hydrolyzed with LiOH to give the corresponding lithium carboxylate. The product was then treated with $KHSO_4$ to liberate the free acid **12**. The other coupling partner, the free amine **15**, was obtained by first benzylating the carboxyl group of Z-Lys(Boc)-OH (**13**), then treating the resulting compound **14** with TFA followed by $NaHCO_3$. Acid **12** and amine **15** were subsequently coupled using benzo-(triazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate (BOP) and NMM to give the protected target **16**. The benzyloxycarbonyl (Z) and benzyl (Bn) protecting groups were then removed by hydrogenation to give the desired compound **7** in quantitative yield. The protecting groups, Bn and Z, were chosen because, unlike *tert*-butoxycarbonyl (Boc) or *tert*-butyl (*t*Bu) protecting groups used in the syntheses of other pyrrolysine analogs (see, for example, Chapters 2.2 and 2.3), these groups can be removed by hydrogenation as opposed to treatment with an acid. This was important because the acetal group of the title compound **7** was expected to be acid sensitive.



***Scheme 1**. Synthesis of pyrrolysine analog **7**.*

The NMR studies showed that no change to the structure of the pyrrolysine analog **7** occurred in any of the prepared solutions listed below.

(1) 25 mM CH_3CO_2H buffer (pH 5.2);

(2) 50 mM Na_2CO_3 (pH 11.2);

(3) 0.5 M NaOAc buffer (pH 4.1);

(4) 0.5 M Na_2CO_3 (pH 11.4);

(5) 1 M TFA (pH 0.5);

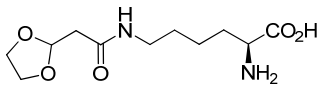
(6) 0.5 TFA (pH 0.5);

(7) 1:1 TFA:D_2O;

(8) 1:1 TFA:D_2O spiked with ethylene glycol;

(9) 0.5 M HCl.

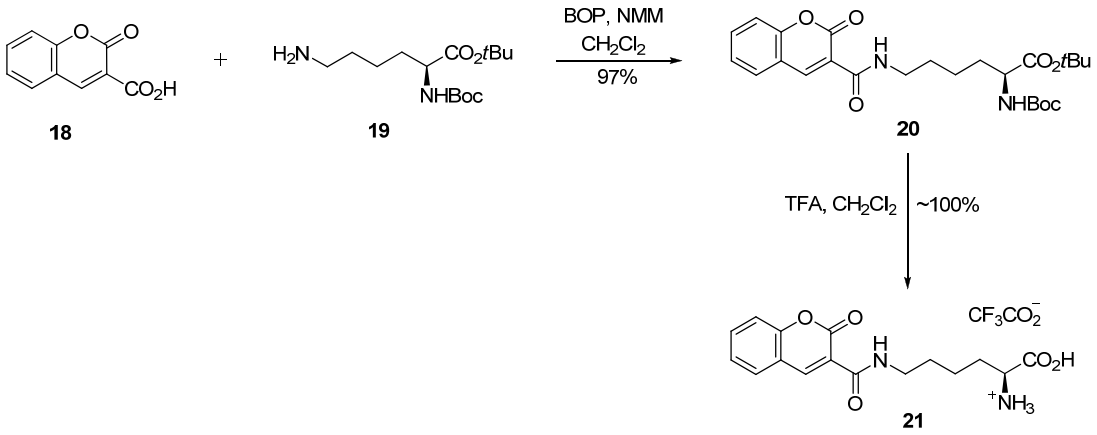
Pyrrolysine analog **7** was synthesized and then used in biochemical studies. The compound is remarkably stable, even under harshly acidic conditions. Attempts to convert the acetal form of the analog into the aldehyde form are ongoing. Future attempts to cleave the acetal may include treating the compound with a Lewis acid, such as a lanthanide. Another possibility would be to make the acetal less stable by inserting a CH_2 group between the carbonyl carbon and the acetal group, giving the compound shown below.



17

2.2. Taggable Pyrrolysine Analog (21)

The taggable pyrrolysine analog **21**, N^6-(2-oxo-2*H*-chromene-3-oyl)lysine trifluoroacetate, was chosen as a target compound because once incorporated into a protein system, the coumarin group would fluoresce and could therefore serve as a tag. Pyrrolysine analog **21** was synthesized with an overall yield of 97% as outlined in Scheme 2, and was subsequently used in biochemical studies.

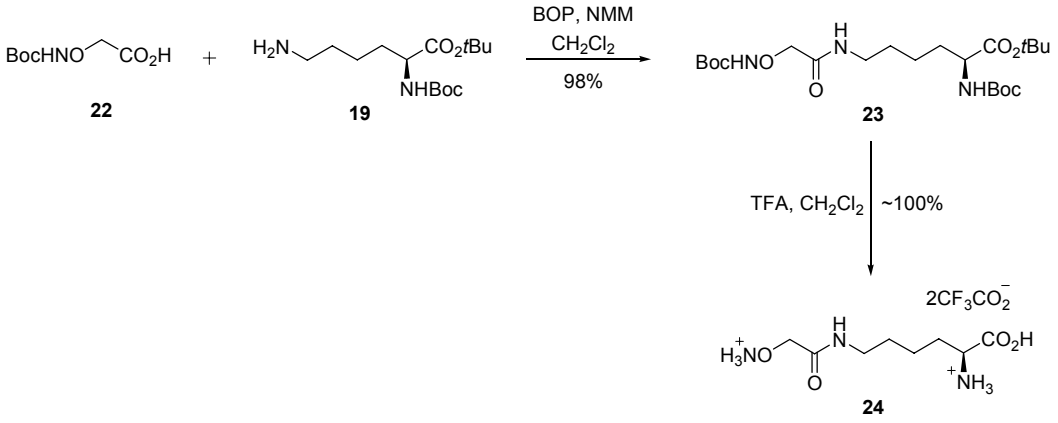


***Scheme 2**. Synthesis of pyrrolysine analog **21**.*

2.3. Taggable Pyrrolysine Analog (24)

The taggable pyrrolysine analog **24**, N^6-(aminoxyacetoyl)lysine ditrifluoroacetate, was chosen as yet another target compound after attempts to cleave the acetal in pyrrolysine analog **7** were unsuccessful. If this new analog **24** could read through the UAG codon, it could then be reacted with a fluorescent aldehyde-containing dye. This would achieve a similar connectivity to that

which would be obtained if **7** was reacted with a fluorescent hydroxylamine-based dye. However, this approach is not preferable because an aldehyde-containing dye would have to be used in excess. Although it would react preferentially with the strongly nucleophilic protein-incorporated hydroxylamine-lysine, it would also react with any free NH_2 groups of the protein. A potential way to avoid these unwanted reactions would be to mop up the excess dye with a hydroxylamine derivative. Pyrrolysine analog **24** was synthesized with an overall yield of 98% as outlined in Scheme 3, and was subsequently used in biochemical studies.



***Scheme 3**. Synthesis of pyrrolysine analog **24**.*

3. Results, Discussion, and Conclusions: Biochemical Studies

3.1. Incorporation into mCherry Protein

The "cherry red" color of the pelleted cells indicated that the taggable pyrrolysine analog **7** was incorporated into the full-length mCherry proteins.

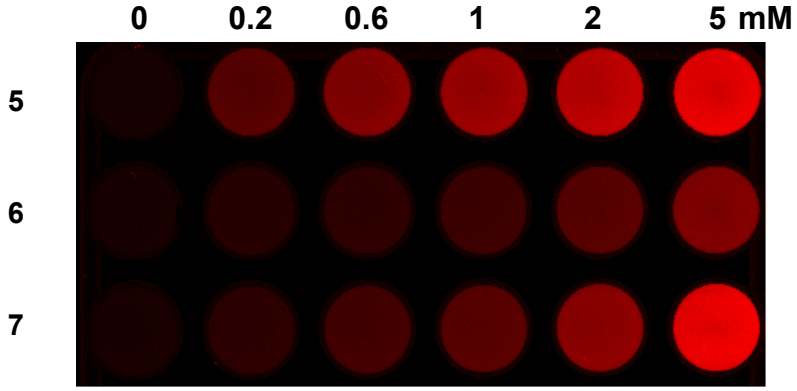


***Figure 2**. Incorporation of pyrrolysine analogs **5-7** into mCherry proteins.*

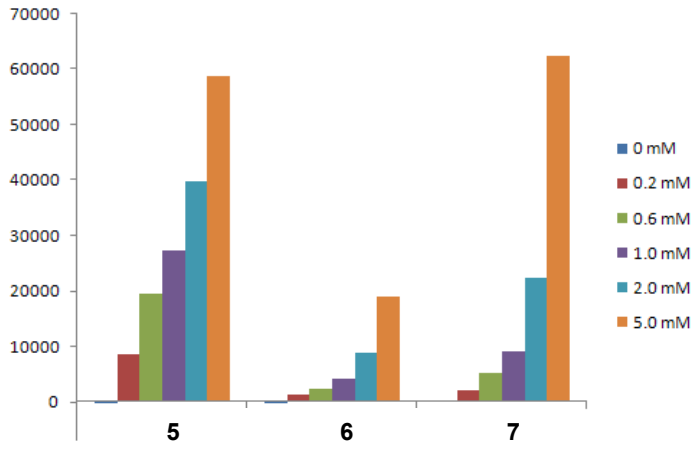


***Figure 3**. Integrated fluorescence count of analogs **5-7** in mCherry proteins.*

Figures 2 and 3 show the readthrough results for **7** in comparison with two previously synthesized pyrrolysine analogs **5** and **6**. The results indicate that **7** reads through the UAG codon more efficiently than **6** at any concentration. In addition, **7** has a comparable read through to **5** at concentrations of 0.2–2 mM, and has a higher read through efficiency than **5** at 5 mM.

Biochemical studies showed that pyrrolysine analogs **21** and **24** do not read through the UAG codon.

3.2. Mass Spectrometry of α-Synuclein Incorporated with Analog (7)

The incorporation of pyrrolysine analog **7** into a full-length α-synuclein protein was verified by a prepared 15% SDS-PAGE gel. Figure 4 shows α-synuclein with **7** at position 44 in lanes 1 and 2. Lane 3 in this figure is blank. Lane 4 is the protein marker (bio-rad precision plus). The results show that α-synuclein with **7** at position 44 runs between the 20 kD and 25 kD bands.

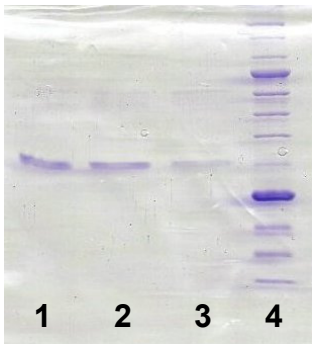


Figure 4*. 15% SDS-PAGE gel of α-synuclein with* ***7*** *at position 44.*

The full length sequence for α-synuclein is as follows: [1]MDVFMKGLSKAKEGVVAAAEK TKQGVAEAAGKTKEGVLYVGSK**TKEGVVHGVATVAEK**TKEQVTNVGGAVVTGVTA VAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQ DYEPEA[140].

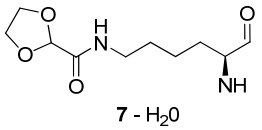
The taggable pyrrolysine analog **7** was incorporated into α-synuclein at position 44 by substituting it for threonine. In the mass spectrometry studies, one tryptic peptide observed at 826.4573^{2+} was identified as **^{44}(7)KEGVVHGVATVAEK^{58}**. The theoretical *m*/*z* value was 826.4543^{2+}, and the mass error was 3.63 ppm. A detailed MS/MS map of **^{44}(7)KEGVVHGVATVAEK^{58}** is shown in Table 1.

Δm ($b^n - b^{n-1}$)	**Measured *m*/*z***	**Sequence**	**Measured *m*/*z***	**Δm ($y^n - y^{n-1}$)**
		(7)		
129.11	486.30	Glu, **E**	1295.49	128.96
57.00	543.30	Gly, **G**	1166.53	57.06
99.01	642.31	Val, **V**	1109.47	99.00
99.05	741.36	Val, **V**	1010.47	99.22
137.13	878.49	His, **H**	911.36	193.95
56.99	935.48	Gly, **G**		
98.97	1034.45	Val, **V**	717.41	99.00
71.06	1105.51	Ala, **A**	618.41	71.14
100.93	1206.44	Thr, **T**	547.27	101.00

99.15	1305.59	Val, **V**	446.27	99.09
70.92	1376.51	Ala, **A**	347.18	70.98
129.08	1505.59	Glu, **E**	276.20	
		Lys, **K**		

***Table 1**. MS/MS map of [44](7)KEGVVHGVATVAEK[58].*

The results indicate that **7** was incorporated into α-synuclein at position 44 in the acetal form of the analog shown below. This form is used for mass calculations because the CO–NH bond is the most common MS cleavage site in peptides.



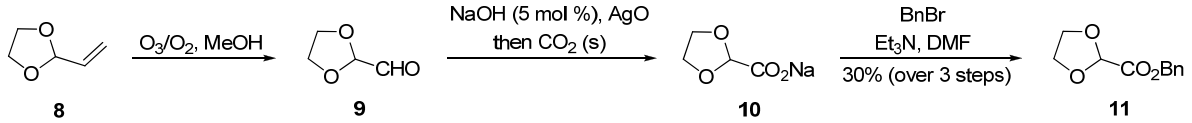
7 - H_2O

Chemical Formula: $C_{10}H_{17}N_2O_4$
Exact Mass: 229.12
Molecular Weight: 229.25
m/z: 229.12 (100.0%), 230.12 (11.7%)

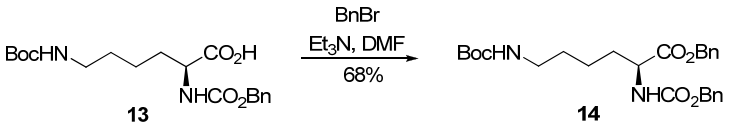
If a method to convert the acetal form of analog **7** into the aldehyde form can be found, future biochemical studies will be performed with the protein-incorporated analog. Such studies would include analyzing the properties of salt bridges between neighboring lysine and glutamate residues in certain proteins, such as α-synuclein, by site-specifically substituting glutamate residues with the aldehyde form of the analog **7**.

4. Experimental Methods: Organic Chemistry Studies

4.1. Synthesis of Taggable Pyrrolysine Analog (7)



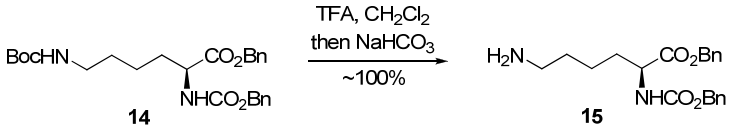
Benzyl 1,3-dioxolane-2-carboxylate (11): Following a published procedure for the synthesis of methyl 1,3-dioxolane-2-carboxylate,[6] excess ozone was bubbled through a solution of 2-vinyl-1,3-dioxolane (**8**, 4.7 g, 47 mmol) in methanol (55 mL) at –15 °C for approximately 1 h. The resulting solution was added to aqueous 5% NaOH (83 mL) and silver oxide (23.2 g, 99.5 mmol). After 1 h of rapid stirring, the solution was filtered and the filtrate was washed with ether (100 mL). Excess solid CO_2 was added into the aqueous layer, which was then evaporated under high vacuum to remove the volatiles. The resulting white solid was suspended in hot methanol, the suspension was filtered, and the filtrate was evaporated to dryness. The resulting solid, benzyl bromide (2.85 g, 16.7 mmol), triethylamine (1.35 g, 13.3 mmol), and DMF (40 mL) were combined and stirred for 24 h. The reaction mixture was evaporated under high vacuum. The residue was dissolved in CH_2Cl_2 (100 mL), washed successively with water and brine, and dried ($MgSO_4$). The suspension was filtered, and the filtrate was evaporated. Purification by flash chromatography (silica gel; hexane/EtOAc, 9/1 → 5/1) gave ester **11** (5.9 g, 30%) as a white solid: R_f = 0.60 (hexanes/EtOAc, 1/1). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.95–4.17 (m, 4H), 5.19 (s, 2H), 5.38 (s, 1H), and 7.29–7.45 (m, 5H) ppm. $^{13}C\{^1H\}/^{13}C$ DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 65.3 (*C*H_2), 66.8 (*C*H_2), 98.3 (*C*H), 128.0 (*C*H), 128.3 (*C*H), 128.4 (*C*H), 135.0 (ipso *C*), and 168.1 (ipso *C*) ppm. IR ($CHCl_3$). ν_{max} = 3019, 3028, 1751, 1282, 1228, 1224, 1221, 1209, 1205, 1127, and 1037 cm^{-1}.



Benzyl (–)-N^2-benzyloxycarbonyl-N^6-(*tert*-butoxycarbonyl)lysinate (14): Acid **13** (10.0 g, 26.3 mmol) was dissolved in DMF (100 mL) in the dark in a round-bottomed flask with an N_2 inlet. Triethylamine (31.6 g, 0.73 mmol) and benzyl bromide (39.5, 1.45 mmol) were added to

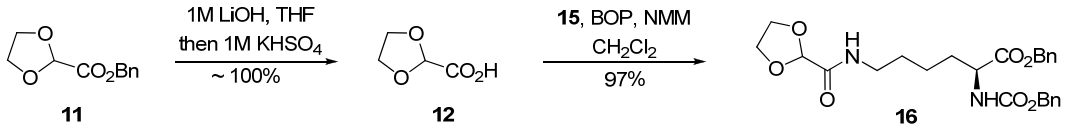
[6] P. D. Dalsin, J. Hine, *J. Am. Chem. Soc.* **1972**, *94*, 6998–7002.

the solution and the reaction mixture was stirred for 24 h, evaporated under high vacuum, and dissolved in CH_2Cl_2 (250 mL). The solution was washed successively with water and brine, dried ($MgSO_4$), filtered, and evaporated. Purification by flash chromatography (silica gel; CH_2Cl_2/EtOAc, 9/1) gave ester **14** (12.4 g, 68%) as a white solid: $[\alpha]^{20}_D$ –5.2 (*c* 0.9, $CHCl_3$). R_f = 0.40 (CH_2Cl_2/EtOAc, 9/1). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.34–1.66 (m, 13H), 1.72–2.06 (m, 2H), 2.93–3.36 (m, 2H), 4.22–4.72 (m, 1H), 4.78–5.09 (m, 1H), 5.12–5.42 (m, 4H), 5.66–6.15 (m, 1H), and 7.28–7.76 (m, 10H) ppm. ^{13}C{^{1}H}/^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 22.1 (*C*H_2), 28.1 (*C*H), 29.2 (*C*H_2), 31.6 (*C*H_2), 39.7 (*C*H_2), 53.6 (*C*H_3), 53.8 (*C*H_3), 53.8 (*C*H_3), 66.6 (*C*H_2), 66.7 (*C*H_2), 78.6 (ipso *C*), 127.6 (*C*H), 127.8 (*C*H), 128.0 (*C*H), 128.1 (*C*H), 128.2 (*C*H), 128.3 (*C*H), 135.1 (ipso *C*), 136.0 (ipso *C*), 155.3 (ipso *C*), 155.8 (ipso *C*), and 172.1 (ipso *C*) ppm. IR ($CHCl_3$). ν_{max} = 1713, 1509, 1240, 1217, 1209, and 1172 cm^{-1}.



14 **15**

Benzyl N^2-(benzyloxycarbonyl)lysinate (15): A solution of amine **14** (2.36 g, 5.00 mmol) in TFA (5.0 mL), was stirred at room temperature for 1 h. The solution was neutralized with satd $NaHCO_3$, and extracted with CH_2Cl_2. The combined extracts were dried ($MgSO_4$), filtered, and evaporated to give the title compound **15** (1.85 g, ~ 100%) as a clear oil that was used immediately in the following reaction without any further purification. Amine **15**: ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.32–1.62 (m, 6H), 2.65–2.95 (m, 2H), 3.54–4.41 (m, 2H), 4.42–4.62 (m, 1H), 5.10–5.38 (m, 4H), 5.86 (s, 1H), and 7.32–7.55 (m, 10H) ppm. ^{13}C{^{1}H}/^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 22.3 (*C*H_2), 31.2 (*C*H_2), 32.9 (*C*H_2), 40.8 (*C*H_2), 53.8 (*C*H), 66.8 (*C*H_2), 66.9 (*C*H_2), 127.9 (*C*H), 128.0 (*C*H), 128.1 (*C*H), 128.3 (*C*H), 128.3 (*C*H), 128.4 (*C*H), 135.2 (ipso *C*), 136.2 (ipso *C*), 155.9 (ipso *C*), and 172.2 (ipso *C*) ppm.

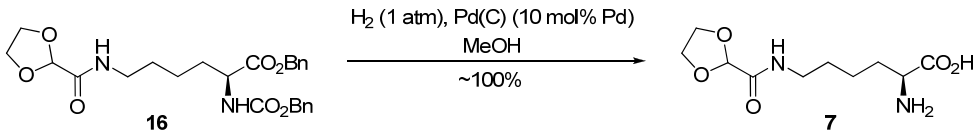


11 **12** **16**

Benzyl (–)-N^2-Benzyloxycarbonyl-N^6-(1,3-dioxolancarbonyl)lysinate (16): A solution of benzyl ester **11** (1.04 g, 5.0 mmol) in THF (30 mL) was treated with 1M LiOH (7.5 mL) and the reaction mixture was stirred for 30 min. 1M $KHSO_4$ (7.5 mL) was added and the solution was

extracted with EtOAc. The combined extracts were dried ($MgSO_4$) and filtered. The filtrate was treated with *N*-methylmorpholine (NMM, 1.1 mL) and evaporated to give the NMM salt of acid **12** that was used immediately without further purification.

Subsequently, to a solution of the NMM salt of the acid **12** in CH_2Cl_2 (50 mL) were added amine **15** (1.85 g, 5.0 mmol), NMM (1.2 mL), and BOP (2.43 g, 5.5 mmol). The reaction mixture was stirred at room temperature for 22 h, washed successively with water and brine, dried ($MgSO_4$), and evaporated in vacuo. Purification by flash chromatography (silica gel; EtOAc/hexanes, 2/1→1/1) gave the protected compound **16** (2.28 g, 97%) as a clear oil: $[\alpha]^{20}_D$ –3.4 (*c* 1.2, $CHCl_3$). R_f = 0.20 (EtOAc/hexanes, 2/1). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.25–1.91 (m, 6H), 3.21 (dd, *J* = 12.8 Hz, 6.3 Hz, 2H), 3.90–4.03 (m, 4H), 4.40 (m, 1H), 5.10 (s, 2H), 5.13 (s, 1H), 5.14–5.25 (ABq, *J* = 12.3 Hz, 2H), 5.38–5.44 (*J* = 7.9 Hz, 1H), 6.51 (s, 1H), and 7.26–7.45 (m, 10H) ppm. $^{13}C\{^1H\}/^{13}C$ DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 22.1 (CH_2), 28.3 (CH_2), 31.4 (CH_2), 36.2 (CH_2), 53.5 (*C*H), 65.0 (CH_2), 66.5 (CH_2), 66.6 (CH_2), 98.9 (*C*H), 127.6 (*C*H), 127.7 (*C*H), 127.8 (*C*H), 128.0 (*C*H), 128.1 (*C*H), 128.2 (*C*H), 135.0 (ipso *C*), 135.9 (ipso *C*), 155.8 (ipso *C*), 167.6 (ipso *C*), and 172.0 (ipso *C*) ppm. IR ($CHCl_3$). ν_{max} = 3012, 1720, 1688, 1512, 1348, 1238, 1229, 1223, 1214, 1208, 1189, and 1122 cm^{-1}.



16 **7**

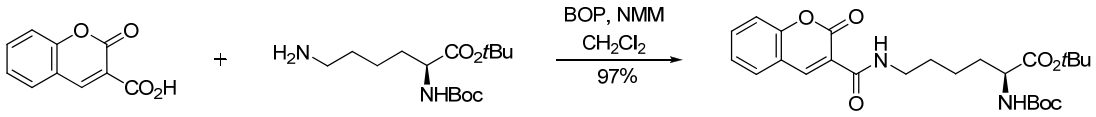
(+)-N^6-(1,3-dioxolancarbonyl)lysine (7): A solution of the protected pyrrolysine analog **16** (341 mg, 0.73 mmol) in MeOH (6 mL) was hydrogenated at normal pressure in the presence of 10% Pd/C (72 mg) for 90 min. Water (10 mL) was added, the suspension was passed through a short column filled with Celite®, and the filtrate was evaporated in vacuo to give the title compound **7** (176 g, 99%) as a white solid that was used in biochemical studies without any further purification. Compound **7**: $[\alpha]^{20}_D$ +1.7 (*c* 1.0, H_2O). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.30−1.53 (m, 2H), 1.54–1.73 (m, 2H), 1.76–2.02 (m, 2H), 3.19–3.40 (t, *J* = 6.9 Hz, 2H), 3.59–3.80 (t, *J* = 5.0 Hz, 1H), 4.07 (s, 4H), and 5.32 (s, 1H) ppm. $^{13}C\{^1H\}/^{13}C$ DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 21.8 (CH_2), 27.9 (CH_2), 30.3 (CH_2), 38.7 (CH_2), 54.7 (*C*H), 65.4 (CH_2), 98.5 (*C*H), 170.0 (ipso *C*), and 175.2 (ipso *C*) ppm.

4.1.1. pH Stability Studies on Taggable Pyrrolysine Analog (7)

(1) 1M acetic acid buffer (0.5 mL, pH 4.9) was added to D_2O (21.5 mL, 23.65 g) to create a 25 mM acetic acid buffer (pH 5.2). (2) Na_2CO_3 (26.5 mg) was added to D_2O (5.0 mL, 5.5 g) resulting in 50 mM sodium carbonate (pH 11.2). To solutions (1) and (2), analog **7** (5.0 mg) was added. ^{1}H NMR studies were performed immediately and again after 24 h. To each NMR sample, lysine monohydrochloride (2.0 mg) was added and ^{1}H NMR studies were performed again.

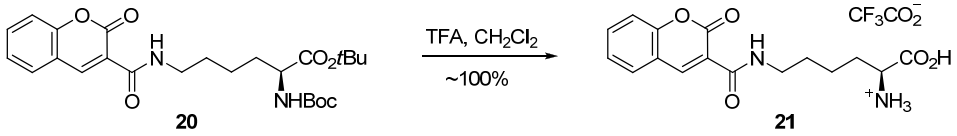
(3) 1 M NaOAc buffer (2.5 mL, pH 4.0) was added to D_2O (2.5 mL, 2.75 g) to create a 0.5 M NaOAc buffer (pH 4.1). (4) Na_2CO_3 (132.5 mg) was added to D_2O (2.5 mL, 2.75 g) to create 0.5 M Na_2CO_3 (pH 11.4). (5) TFA (385 μL) was added to D_2O (5.1 mL, 5.61 g) to create 1M TFA (pH 0.5). (6) TFA (38.5 μL) was added to D_2O (5.0 mL, 5.5 g) to create 0.5 M TFA (pH 0.5). (7) TFA (0.4 mL) was added to D_2O (0.4 mL, 0.44 g) to create a 1:1 solution. (8) TFA (0.4 mL) was added to D_2O (0.4 mL, 0.44 g) to create a 1:1 solution. Ethylene glycol (2 drops) was added to the solution. (9) 12.1 N HCl (80 μL) was added to D_2O (2.0 mL, 2.2 g) to create 0.5 M HCl. To solutions (3), (4), (5), (6), (7), (8), and (9), analog **7** (2.5 mg) was added. ^{1}H NMR studies were performed immediately and again after 24 h. To each NMR sample, lysine monohydrochloride (1.0 mg) was added and proton NMR studies were performed again.

4.2. Synthesis of Taggable Pyrrolysine Analog (21)



***tert*-Butyl (+)-N^2-(*tert*-Butoxycarbonyl)-N^6-(2-oxo-2*H*-chromene-3-oyl)lysinate (20):** To a solution of coumarin **18** (213 mg, 1.12 mmol) in CH_2Cl_2 (5 mL) were added amine **19** (339 mg, 1.12 mmol), NMM (246 μL, 2.24 mmol), and BOP (545 mg, 1.23 mmol). The reaction mixture was stirred at room temperature for 24 h and then partitioned between satd $NaHCO_3$ and CH_2Cl_2. The phases were separated and the extraction was completed with additional portions of CH_2Cl_2. The combined organic extracts were dried ($MgSO_4$) and evaporated in vacuo. Purification by flash chromatography (silica gel; CH_2Cl_2/EtOAc, 3/1) gave the titled compound **20** (0.53 g, 97%)

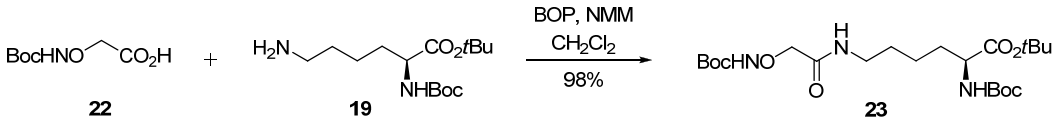
as a white solid: $[\alpha]^{20}_{D}$ +7.5 (*c* 1.1, $CHCl_3$). R_f = 0.40 (CH_2Cl_2/EtOAc, 3/1). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.33–1.79 (m, 24H), 3.32–3.47 (q, *J* = 6.5 Hz, 2H), 4.04–4.16 (m, 1H), 5.02–5.13 (d, *J* = 7.7 Hz, 1H), 7.24–7.38 (m, 2H), 7.54–7.68 (m, 2H), 8.70–8.79 (m, 1H), and 8.83 (s, 1H) ppm. ^{13}C{^{1}H}/^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 22.5 (*C*H_2), 27.8 (*C*H_3), 28.2 (*C*H_3), 29.0 (*C*H_2), 32.3 (*C*H_2), 39.4 (*C*H_2), 53.7 (*C*H), 79.3 (ipso *C*), 81.5 (ipso *C*), 116.4 (*C*H), 118.3 (ipso *C*), 118.5 (ipso *C*), 125.1 (*C*H), 129.6 (*C*H), 133.8 (*C*H), 148.0 (*C*H), 154.2 (ipso *C*), 155.2 (ipso *C*), 161.2 (ipso *C*), 161.3 (ipso *C*), 171.7 (ipso *C*) ppm. IR ($CHCl_3$). ν_{max} = 3013, 1719, 1708, 1611, 1570, 1542, 1501, 1456, 1368, 1244, 1209, and 1160 cm^{-1}.



20 **21**

***N*^{6}-(2-Oxo-2*H*-chromene-3-oyl)lysine trifluoroacetate (21):** The protected pyrrolysine analogue **20** (104 mg, 0.22 mmol) in TFA (2 mL) and CH_2Cl_2 (2 mL) was stirred at room temperature for 1 h, and the resulting solution was evaporated in vacuo to give the title compound **21**·TFA as a white solid (95 mg, ~100%) that was used in the subsequent biochemical studies without any further purification. Acid **21**·TFA: ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.27–1.92 (m, 7H), 3.25–3.42 (m, 2H), 3.84–3.97 (m, 1H), 7.40–7.57 (m, 2H), 7.70–7.81 (t, *J* = 7.2 Hz, 1H), 7.94–8.05 (d, *J* = 7.1 Hz, 1H), 8.19–8.48 (m, 2H), and 8.58–9.04 (m, 2H) ppm. ^{13}C{^{1}H}/^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 21.8 (*C*H_2), 28.4 (*C*H_2), 29.6 (*C*H_2), 38.6 (*C*H_2), 51.8 (*C*H), 116.0 (*C*H), 118.4 (ipso *C*), 118.9 (ipso *C*), 125.1 (*C*H), 130.2 (*C*H), 134.0 (*C*H), 147.3 (*C*H), 153.8 (ipso *C*), 160.3 (ipso *C*), 161.0 (ipso *C*), 171.0 (ipso *C*) ppm.

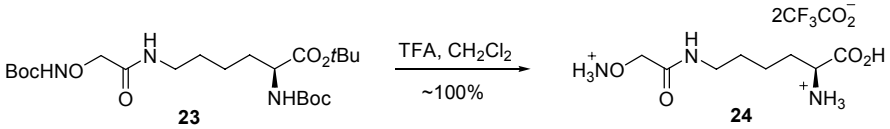
4.3. Synthesis of Taggable Pyrrolysine Analog (24)



22 **19** **23**

***tert*-Butyl (+)-*N*^{2}-(*tert*-Butoxycarbonyl)-*N*^{6}-(*N*-*tert*-butoxycarbonylaminoxyacetoyl)lysinate (23):** To a solution of acid **22** (235 mg, 1.23 mmol) in CH_2Cl_2 (5 mL) were added amine **19** (372 mg, 1.23 mmol), NMM (297 μL, 2.71 mmol), and BOP (598 mg, 1.35 mmol). The reaction mixture was stirred at room temperature for 9 h, washed with satd $NaHCO_3$ and extracted with

CH_2Cl_2. The combined extracts were dried ($MgSO_4$) and evaporated in vacuo. Purification by flash chromatography (silica gel; EtOAc/hexanes, 1/1 → EtOAc) gave the protected compound **23** (575 mg, 98%) as a clear oil: $[\alpha]^{20}_D$ +3.5 (*c* 0.6, $CHCl_3$). R_f = 0.70 (EtOAc). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.26–1.74 (m, 33H), 3.11–3.30 (q, J = 6.5 Hz, 2H), 3.98–4.14 (m, 1H), 4.21 (s, 2H), 4.99–5.14 (d, J = 7.7 Hz, 1H), 8.12 (s, 1H), and 8.46 (s, 1H) ppm. ^{13}C{^{1}H}/^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 22.4 (*C*H_2), 27.8 (*C*H_3), 28.0 (*C*H_3), 28.2 (*C*H_3), 28.8 (*C*H_2), 32.2 (*C*H_2), 38.5 (*C*H_2), 53.7 (*C*H), 75.9 (*C*H_2), 79.4 (ipso *C*), 81.6 (ipso *C*), 82.4 (ipso *C*), 155.3 (ipso *C*), 157.8 (ipso *C*), 168.8 (ipso *C*), and 171.8 (ipso *C*) ppm. IR ($CHCl_3$). ν_{max} = 3025, 3015, 1730, 1709, 1602, 1370, 1252, and 1156 cm^{-1}.



(+)-N^6-(Aminoxyacetoyl)lysine Ditrifluoroacetate (24): The protected pyrrolysine analogue **23** (338 mg, 0.71 mmol) in TFA (5 mL) and CH_2Cl_2 (5 mL) was stirred at room temperature for 2 h, and the resulting solution was evaporated in vacuo to give the title compound **24**·2TFA as a white solid (155 mg, ~100%) that was used in the subsequent biochemical studies without any further purification. Pyrrolysine analog **24**·2TFA: $[\alpha]^{20}_D$ +10.7 (*c* 2.0, MeOH). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 1.45–1.66 (m, 4H), 1.86–2.06 (m, 2H), 3.15–3.45 (m, 3H), 3.90–4.05 (t, J = 6.4 Hz, 1H), 4.46 (s, 2H), and 5.12 (m, 7H) ppm. ^{13}C{^{1}H}/^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$): δ = 21.9 (*C*H_2), 28.4 (*C*H_2), 29.7 (*C*H_2), 38.2 (*C*H_2), 52.5 (*C*H), 71.7 (*C*H_2), 112.3–121.0 (q, J = 292 Hz, ipso *C*), 160.9–162.0 (q, J = 35.4 Hz, ipso *C*), 168.9 (ipso *C*), and 170.5 (ipso *C*) ppm.

5. Experimental Methods: Biochemical Studies

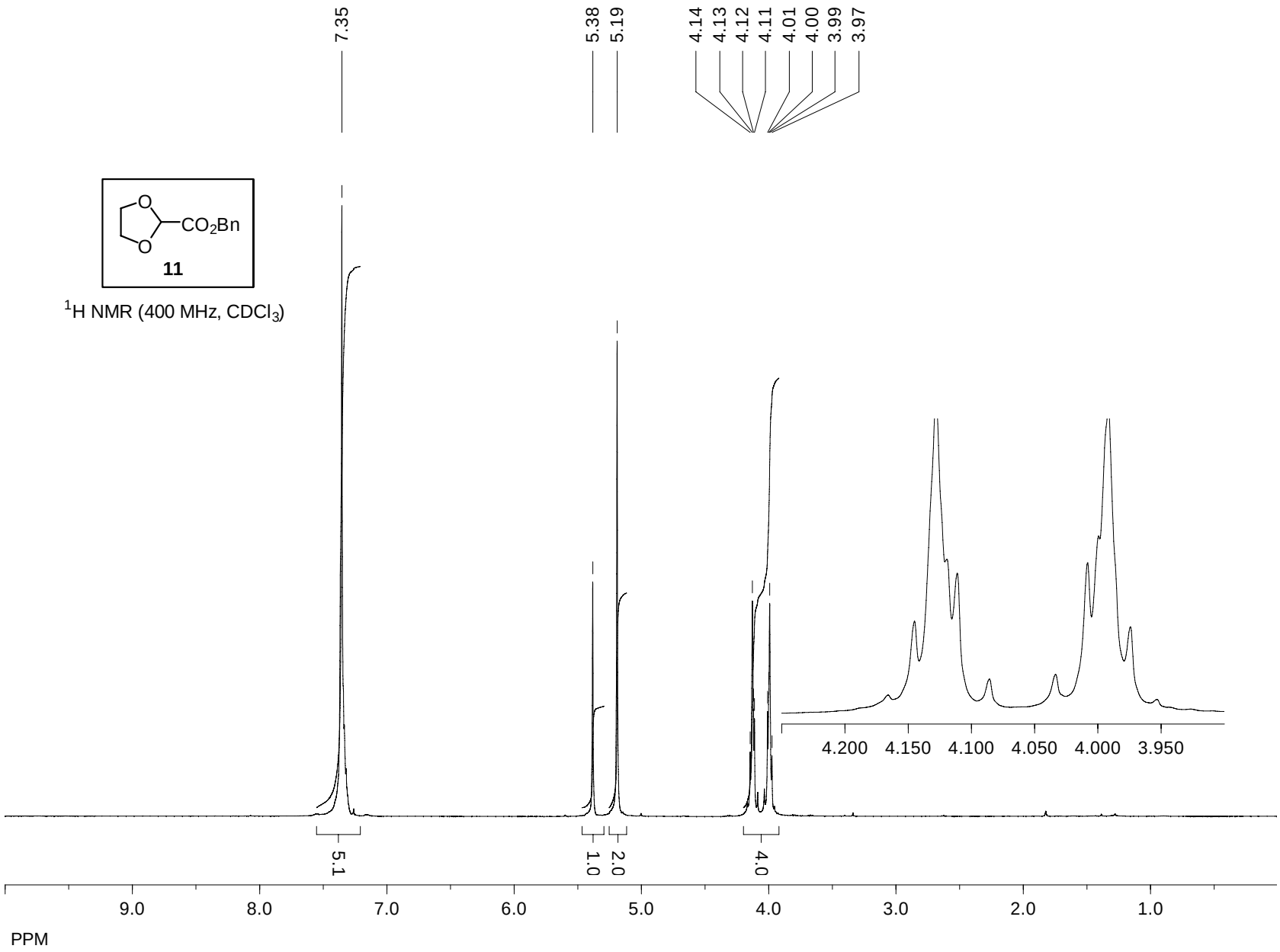
5.1. Incorporation of Pyrrolysine Analogs into mCherry Protein

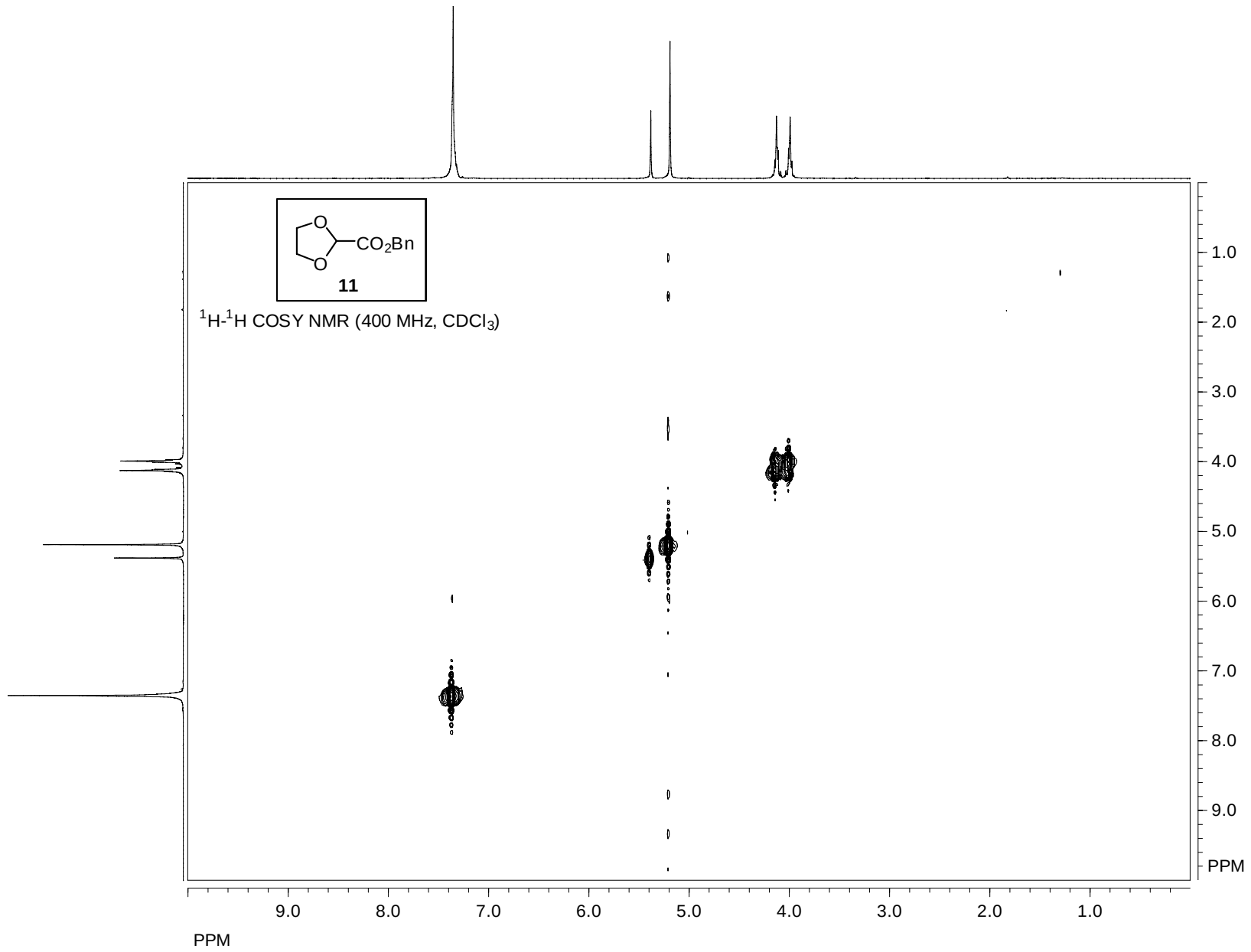
The plasmids encoding the pPylST-mCherry with a UAG codon mutation at position 55 were transformed into *E. coli* BL21(DE3). Cells were grown in Lauria–Bertani (LB) broth containing 100 μg/ml ampicillin at 37 °C. At OD_{600} ~ 1.2, cells were induced with 0.3 mM IPTG and supplemented with 15 mM of the pyrrolysine analog for 3 h to promote the overexpression of mCherry. The cells were then pelleted by centrifugation at 5000 x g for 15 min.

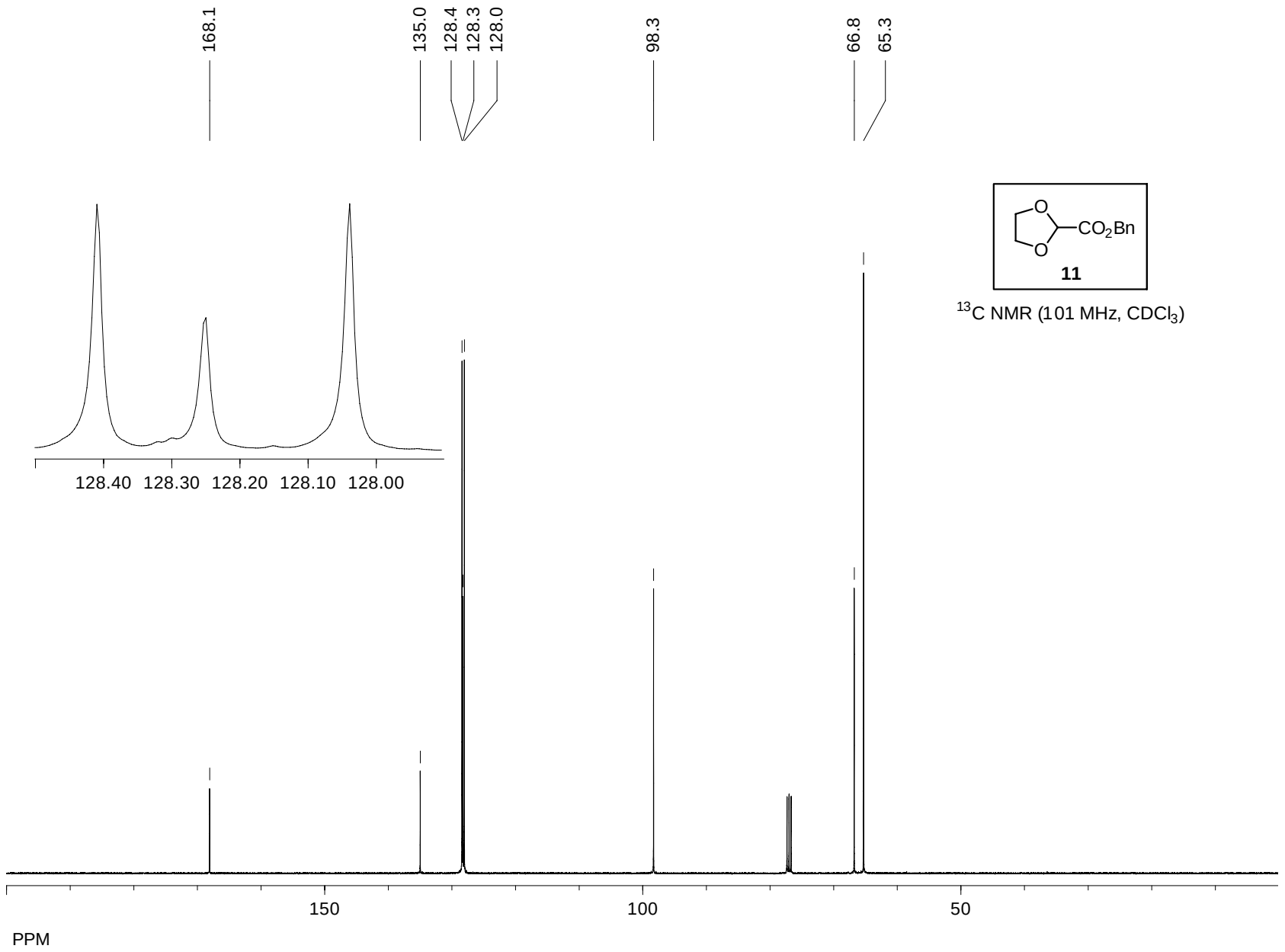
5.2. Incorporation of Analog (7) into α-Synuclein and MS Studies

The plasmids encoding the pPylST-α-synuclein with a UAG codon mutation at position 44 were transformed into *E. coli* BL21(DE3). Cells were grown in 2xYT liquid medium containing 100 μg/ml ampicillin at 37 °C. At OD_{600} ~ 0.6, cells were induced with 0.5 mM IPTG and supplemented with 8 mM of the pyrrolysine analog **7** for 4 h to promote the overexpression of α-synuclein. The cells were pelleted and lysed by sonication. The supernatant lysate was boiled at 90 °C for 10 minutes, followed by centrifugation at 15000 x g for 30 minutes to precipitate unwanted proteins. The supernatant after the boil was loaded onto a 15% SDS-PAGE gel to verify the incorporation of the pyrrolysine analog into the full-length α-synuclein. For mass spectroscopy studies, the α-synuclein band was sliced from gel and digested with chymotrypsin.

Appendix: Copies of ^{1}H and ^{13}C NMR Spectra

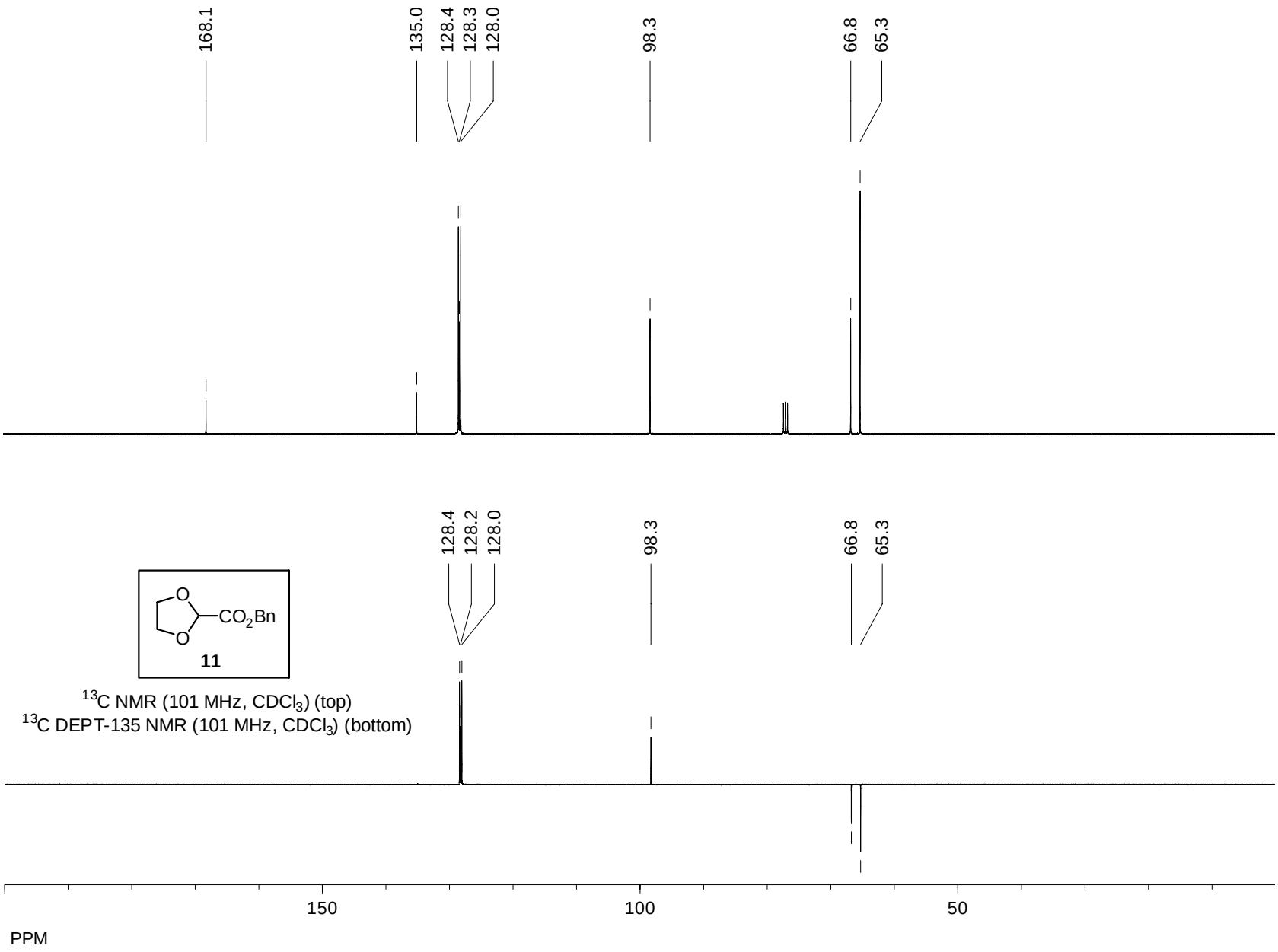

7.35
5.38
5.19
4.14
4.13
4.12
4.11
4.01
4.00
3.99
3.97
CO2Bn
11
1H NMR (400 MHz, CDCl3)
4.200
4.150
4.100
4.050
4.000
3.950
5.1
1.0
2.0
4.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM


CO2Bn
11
1H-1H COSY NMR (400 MHz, CDCl3)
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
PPM
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

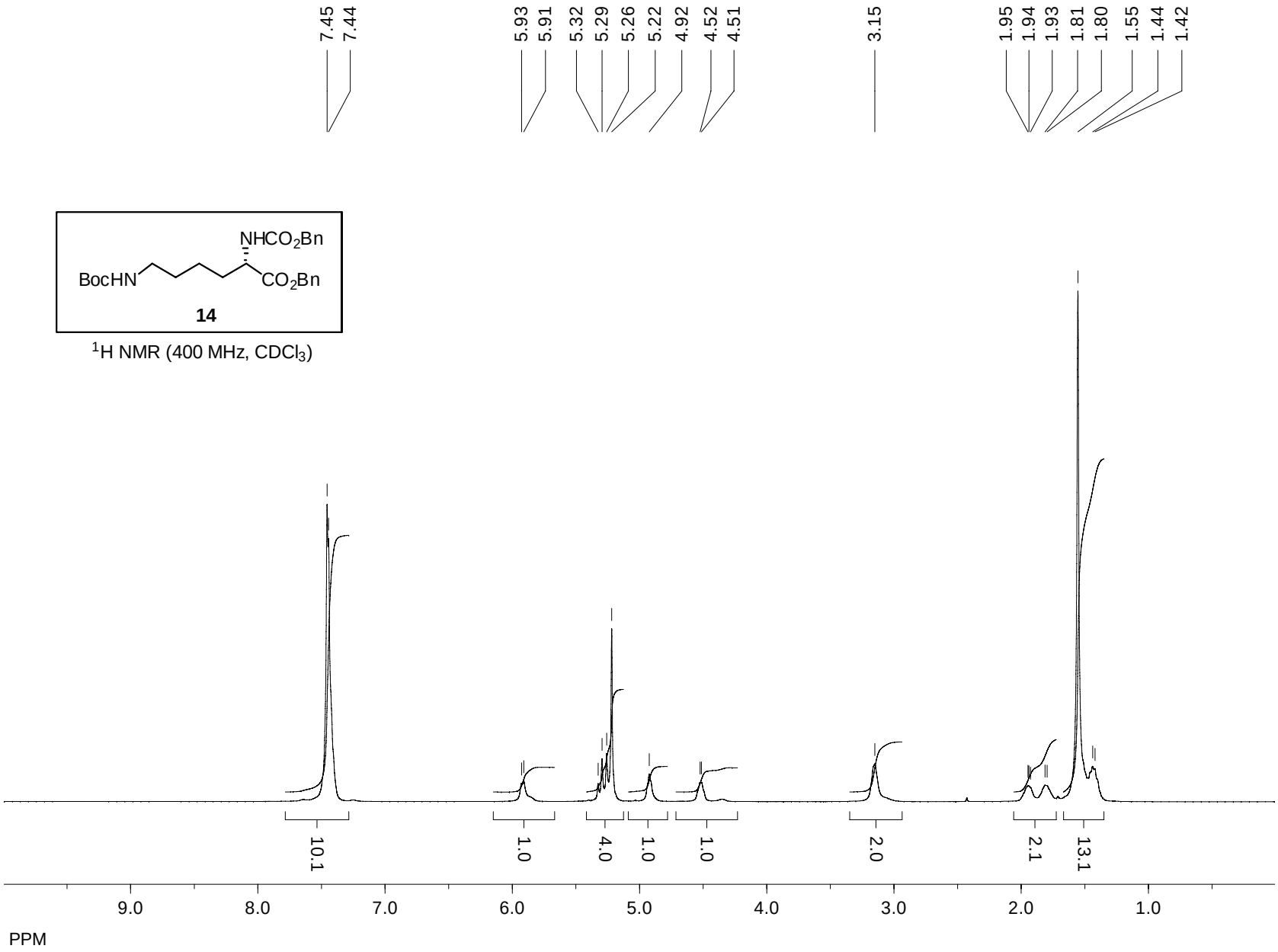


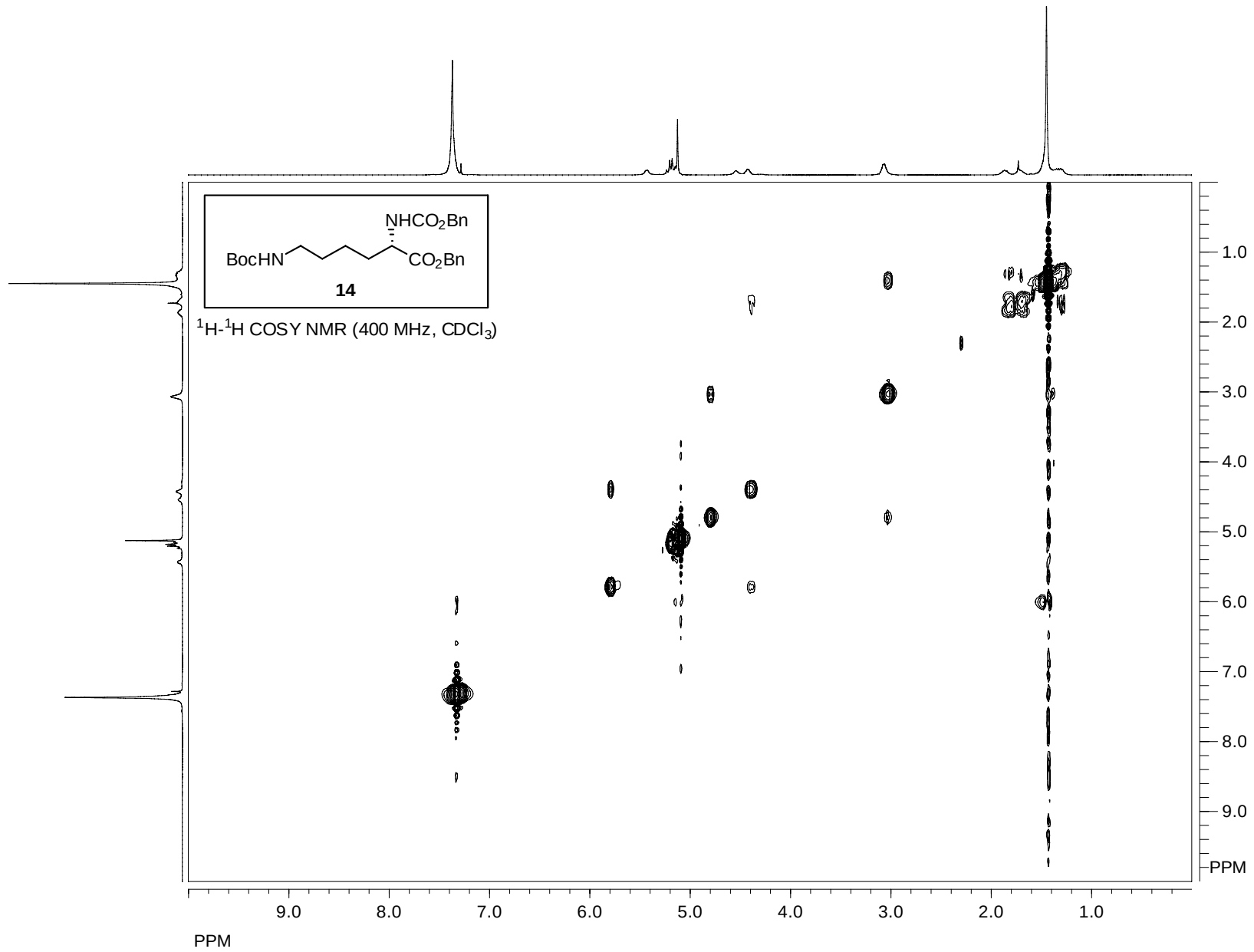
11

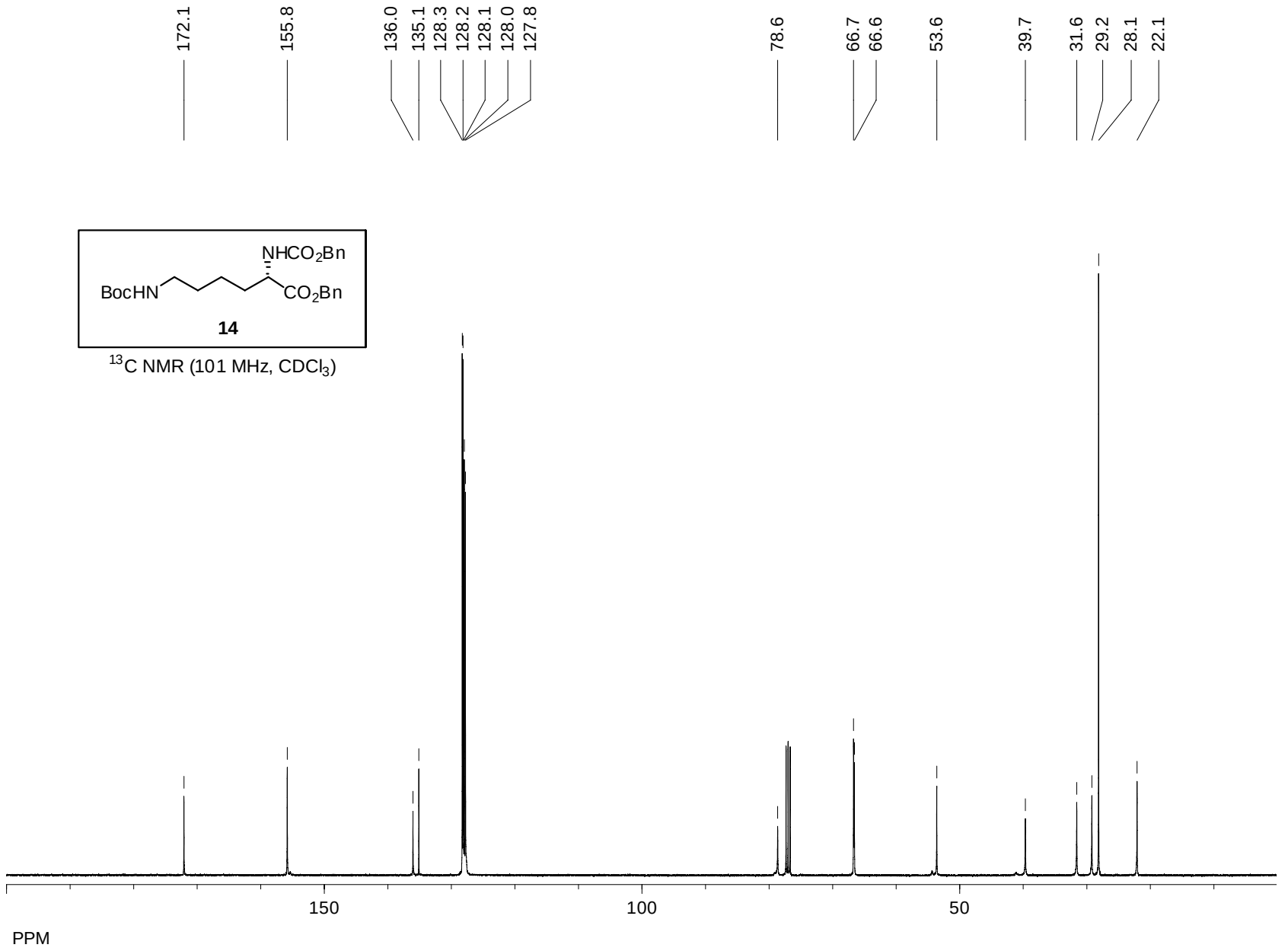
^{13}C NMR (101 MHz, $CDCl_3$)

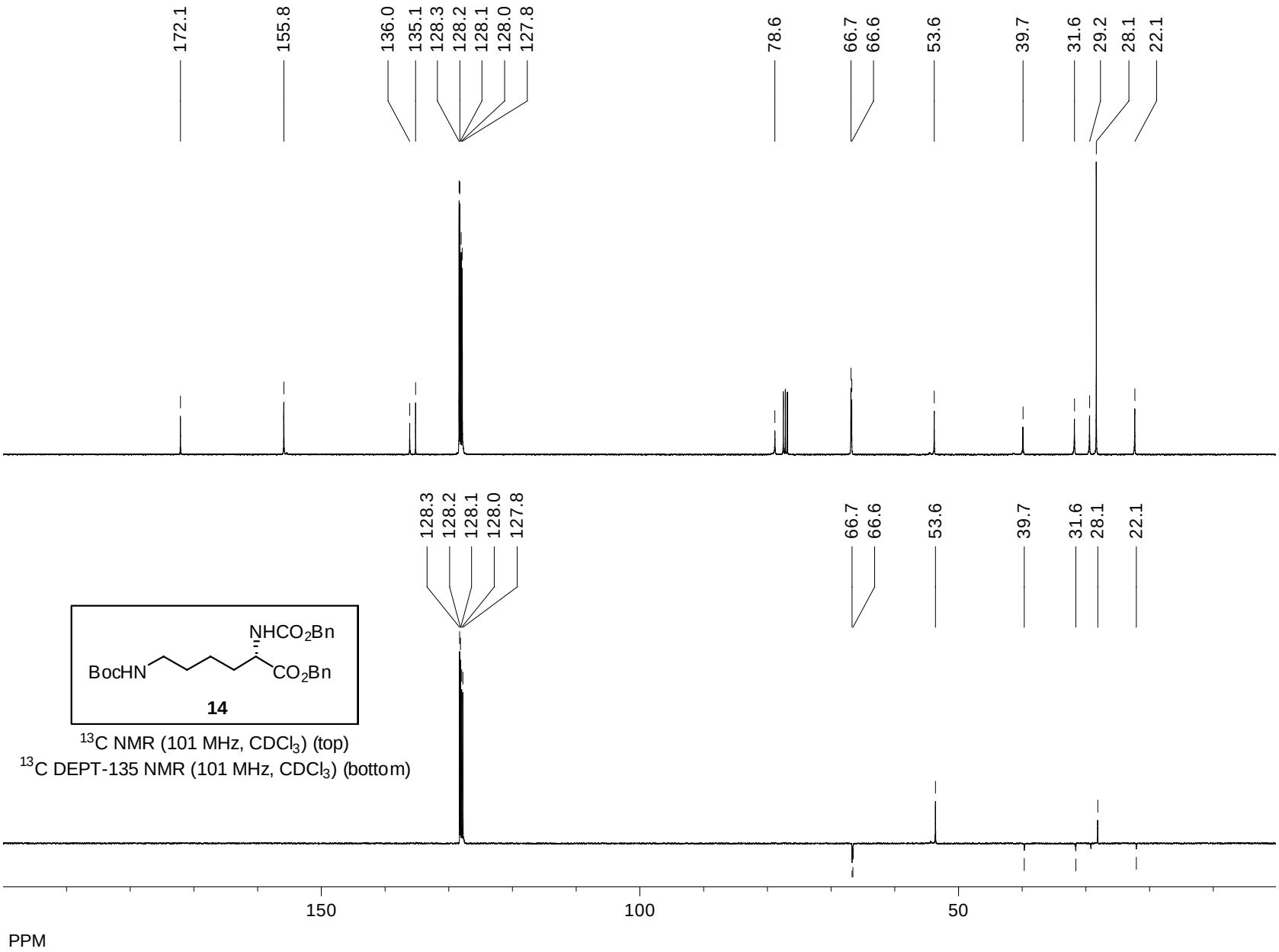


^{13}C NMR (101 MHz, $CDCl_3$) (top)
^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$) (bottom)

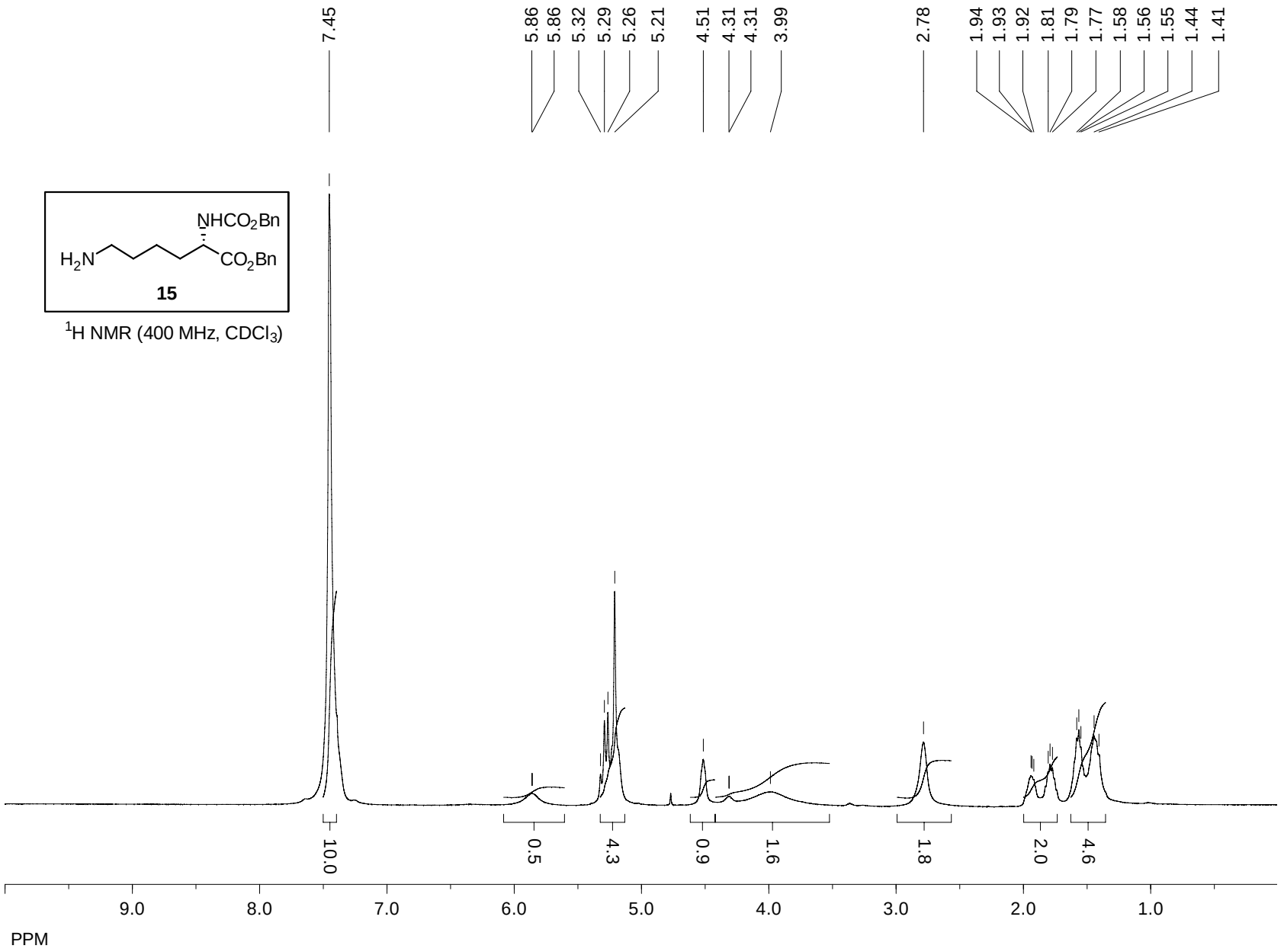

7.45
7.44
5.93
5.91
5.32
5.29
5.26
5.22
4.92
4.52
4.51
3.15
1.95
1.94
1.93
1.81
1.80
1.55
1.44
1.42
NHCO2Bn
BocHN
CO2Bn
14
1H NMR (400 MHz, CDCl3)
10.1
1.0
4.0
1.0
1.0
2.0
2.1
13.1
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

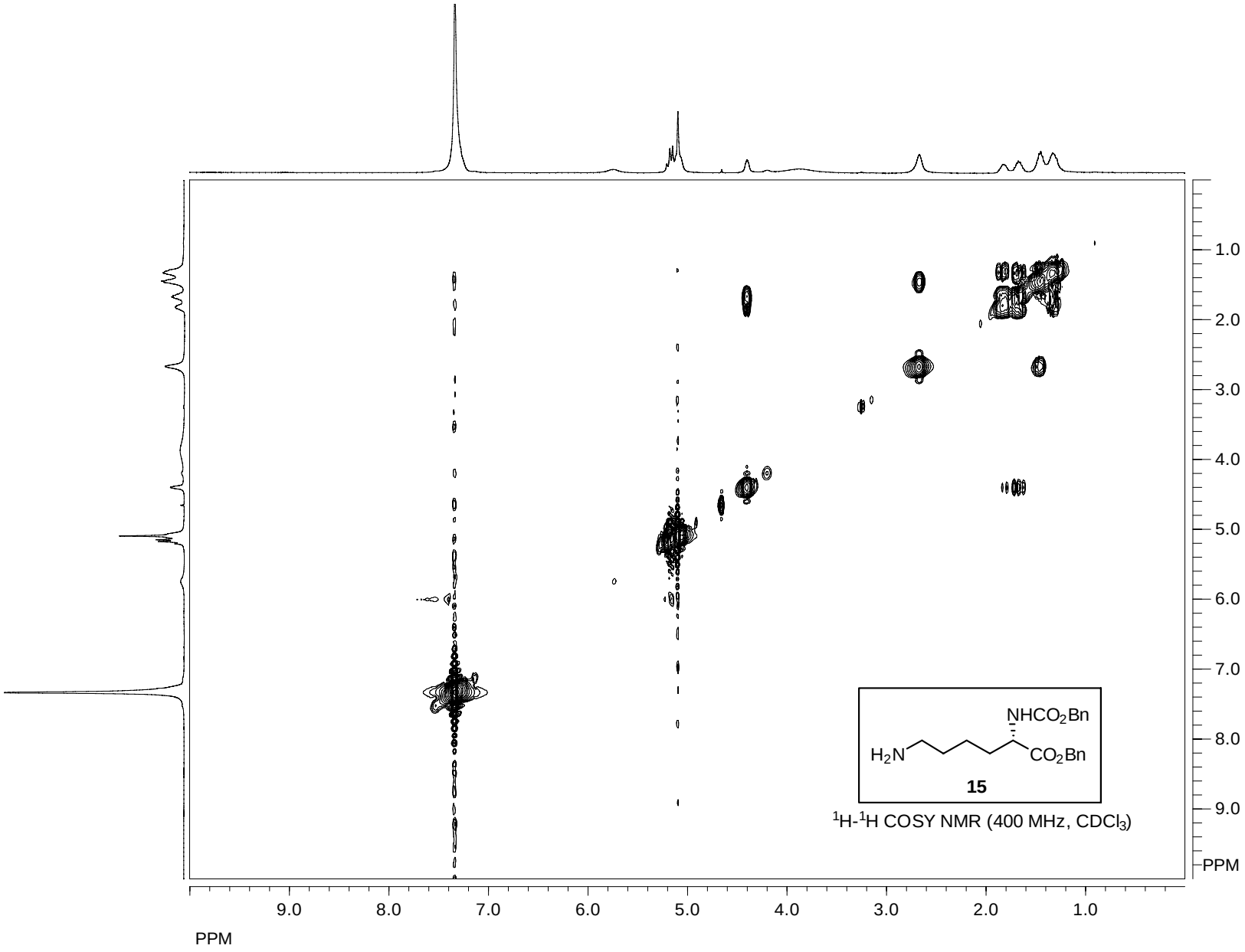

NHCO2Bn
BocHN
CO2Bn
14
1H-1H COSY NMR (400 MHz, CDCl3)
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

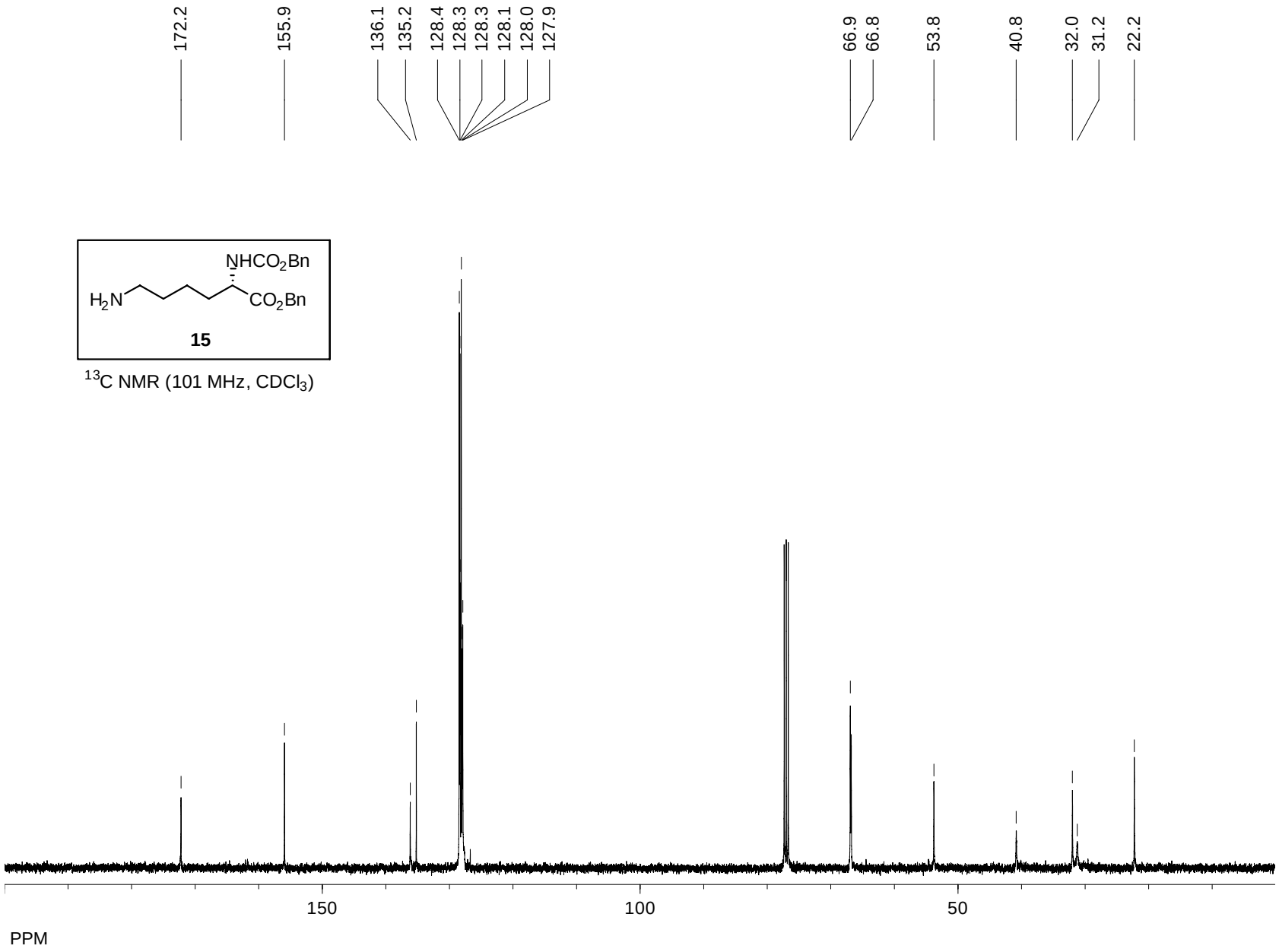

172.1
155.8
136.0
135.1
128.3
128.2
128.1
128.0
127.8
78.6
66.7
66.6
53.6
39.7
31.6
29.2
28.1
22.1
NHCO2Bn
BocHN
CO2Bn
14
13C NMR (101 MHz, CDCl3)
150
100
50
PPM

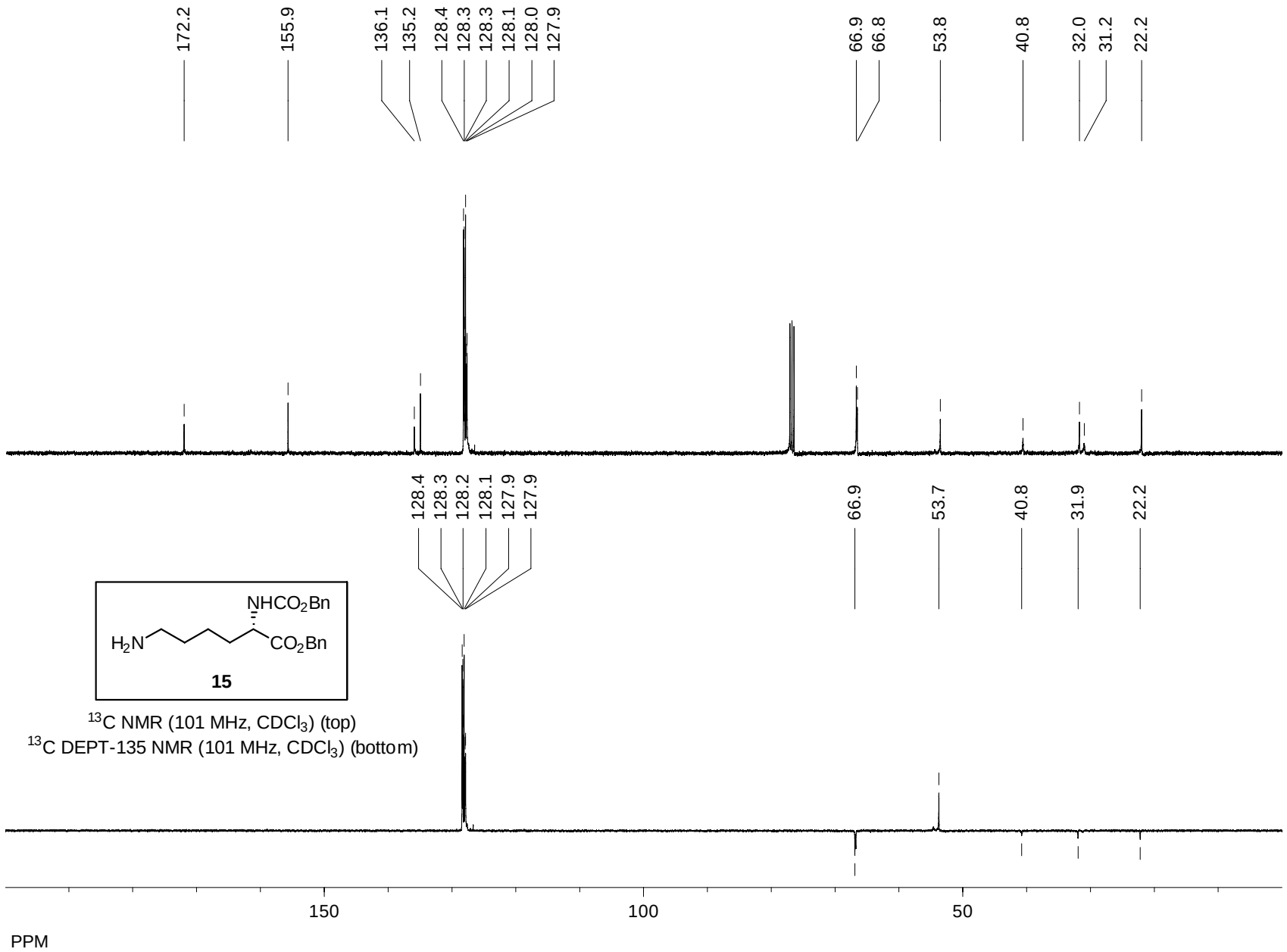


^{13}C NMR (101 MHz, $CDCl_3$) (top)
^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$) (bottom)

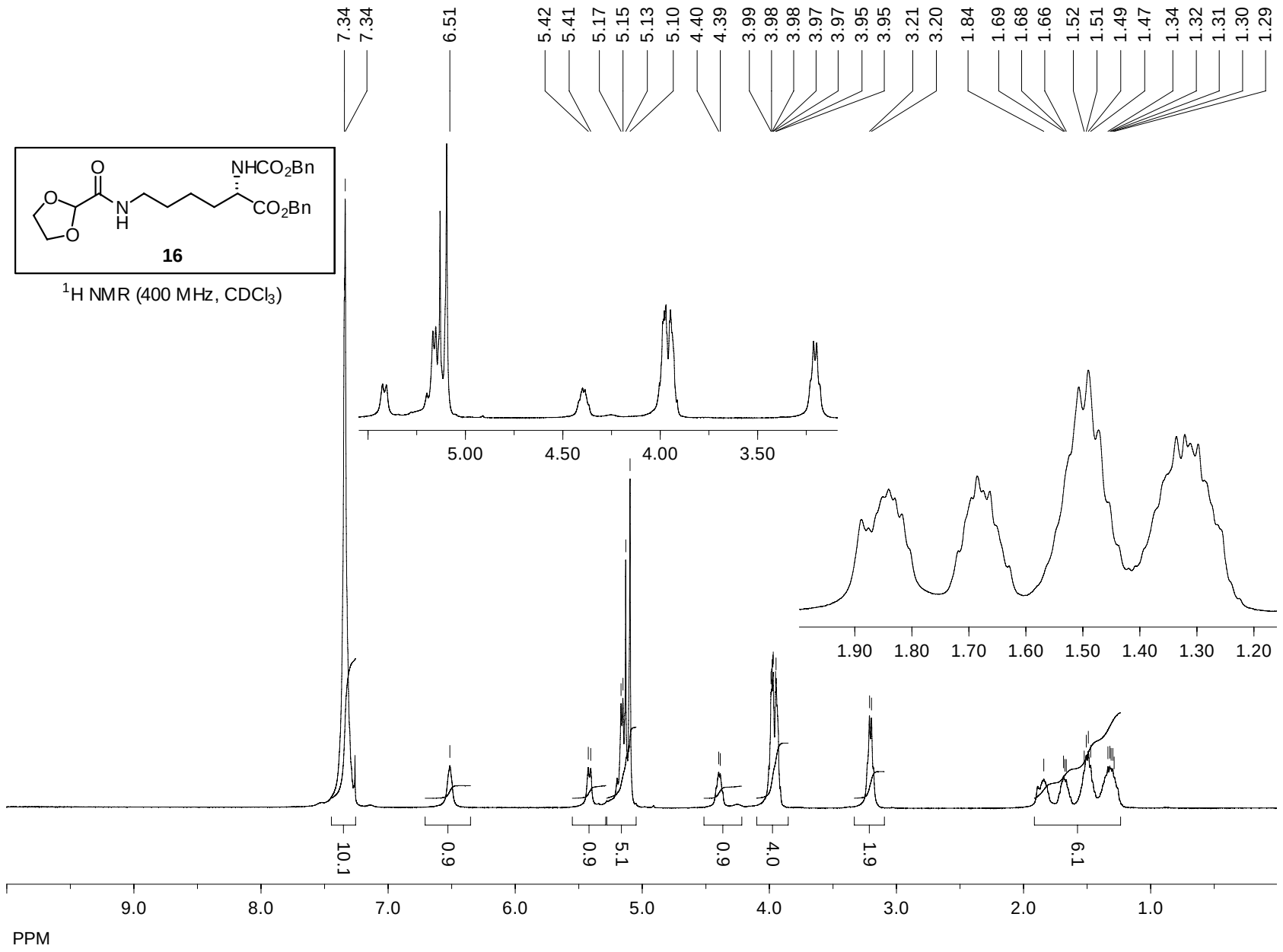

7.45
5.86
5.86
5.32
5.29
5.26
5.21
4.51
4.31
4.31
3.99
2.78
1.94
1.93
1.92
1.81
1.79
1.77
1.58
1.56
1.55
1.44
1.41
NHCO2Bn
H2N
CO2Bn
15
1H NMR (400 MHz, CDCl3)
10.0
0.5
4.3
0.9
1.6
1.8
2.0
4.6
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

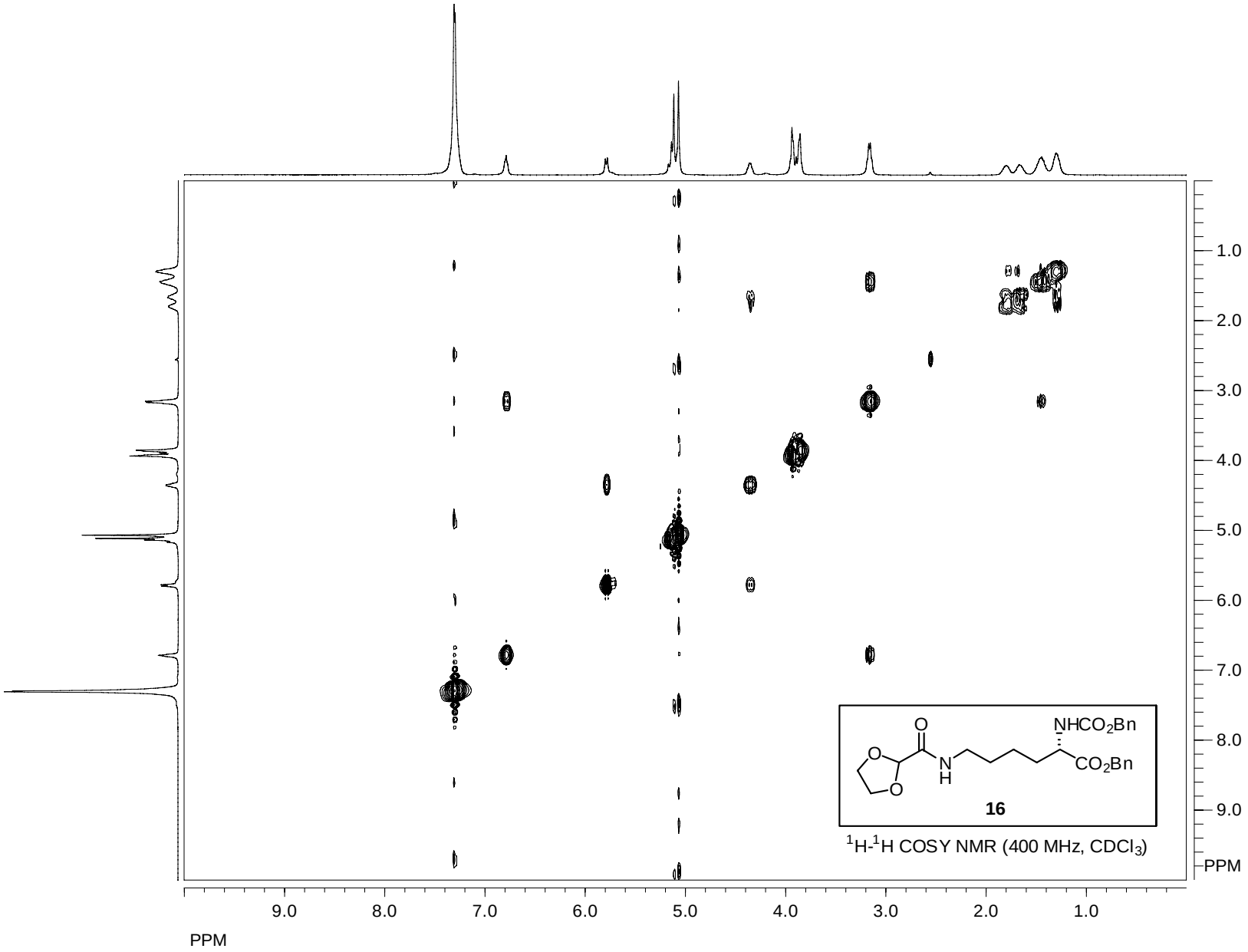

NHCO2Bn
H2N
CO2Bn
15
1H-1H COSY NMR (400 MHz, CDCl3)
PPM
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0


172.2
155.9
136.1
135.2
128.4
128.3
128.3
128.1
128.0
127.9
66.9
66.8
53.8
40.8
32.0
31.2
22.2
NHCO2Bn
H2N
CO2Bn
15
13C NMR (101 MHz, CDCl3)
150
100
50
PPM

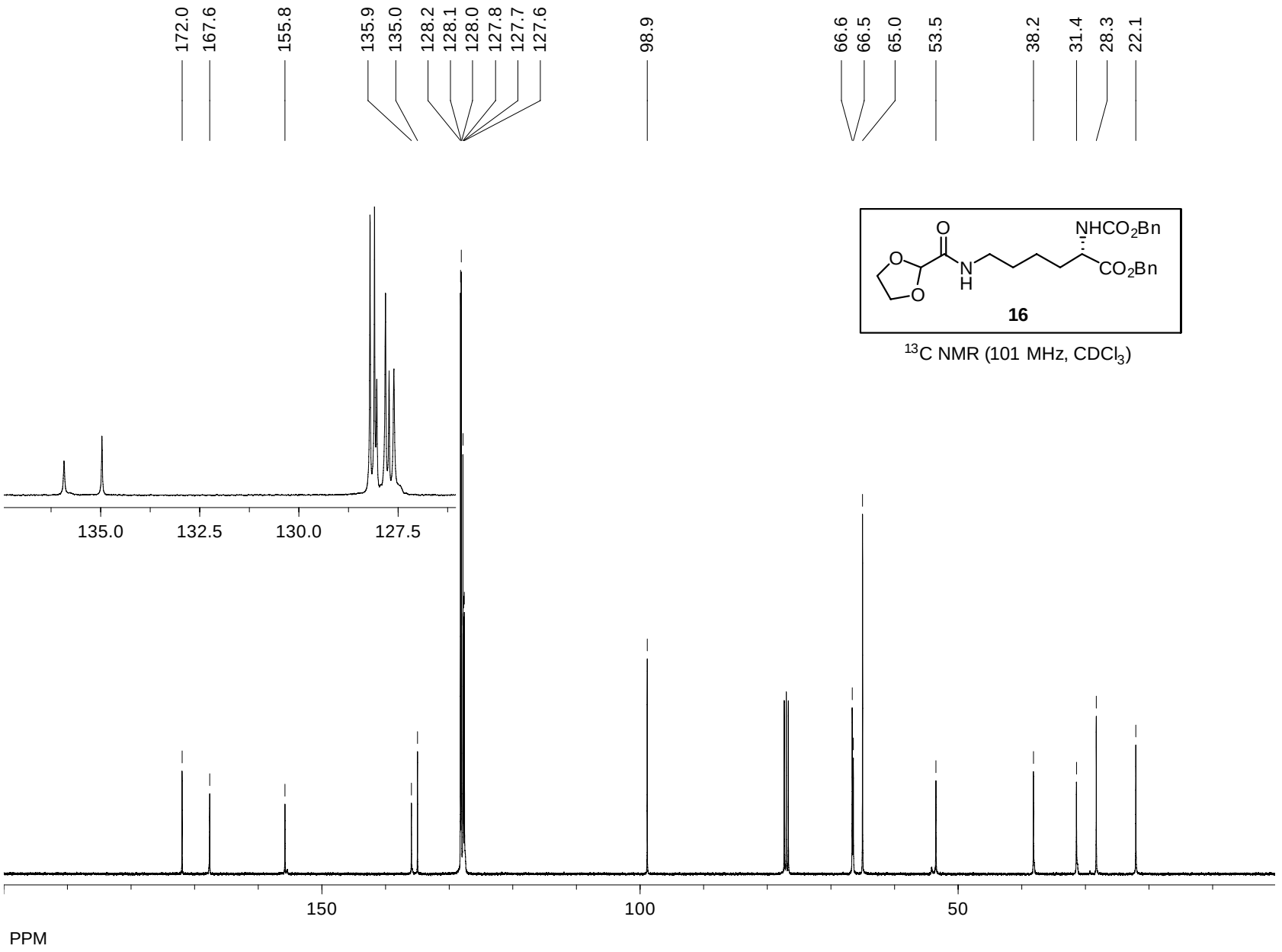


^{13}C NMR (101 MHz, $CDCl_3$) (top)
^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$) (bottom)

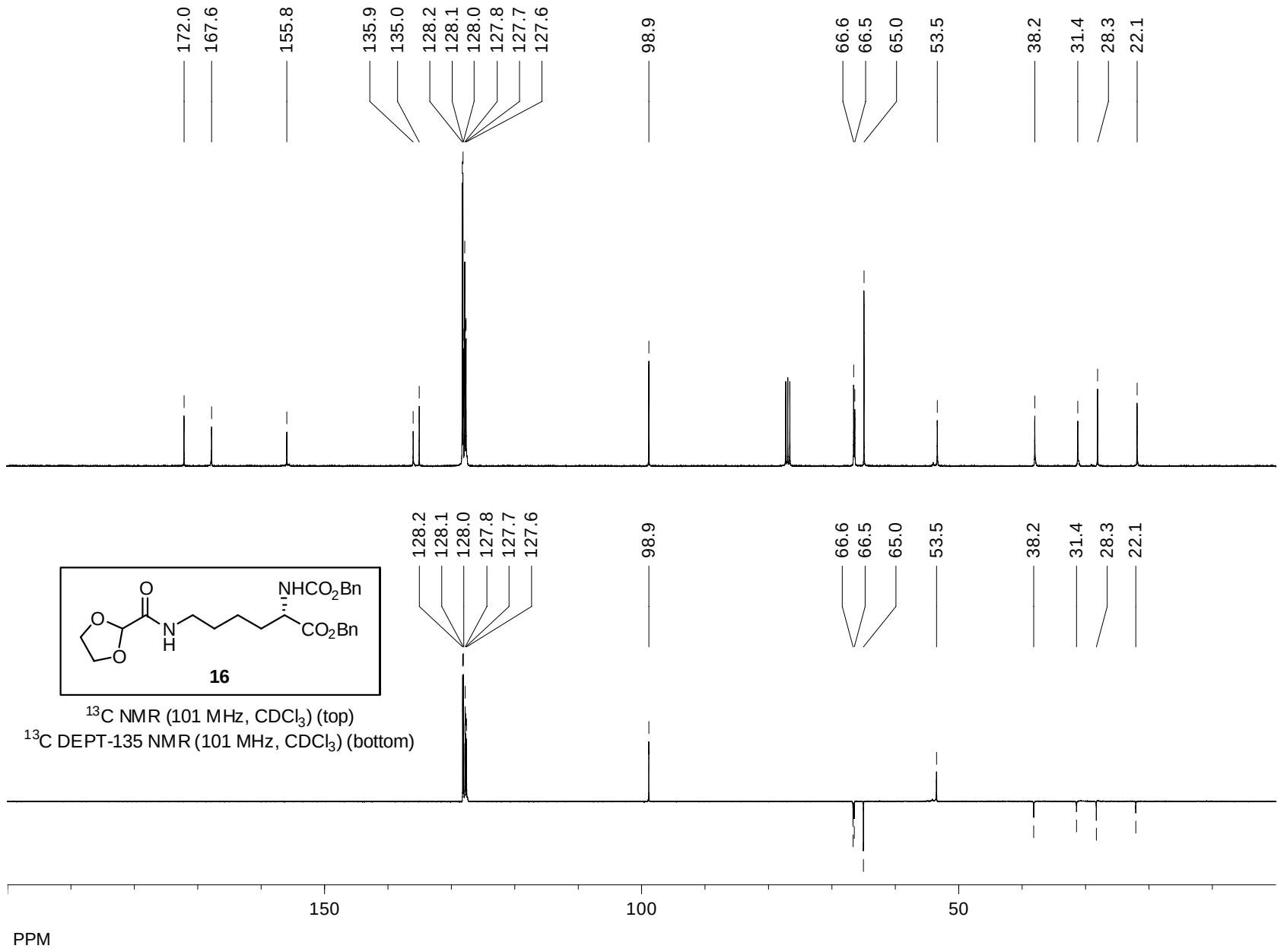

7.34
7.34
6.51
5.42
5.41
5.17
5.15
5.13
5.10
4.40
4.39
3.99
3.98
3.98
3.97
3.97
3.95
3.95
3.21
3.20
1.84
1.69
1.68
1.66
1.52
1.51
1.49
1.47
1.34
1.32
1.31
1.30
1.29
O
NHCO2Bn
O
N
H
CO2Bn
O
16
1H NMR (400 MHz, CDCl3)
5.00
4.50
4.00
3.50
1.90
1.80
1.70
1.60
1.50
1.40
1.30
1.20
10.1
6.9
0.9
5.1
0.9
4.0
1.9
6.1
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM



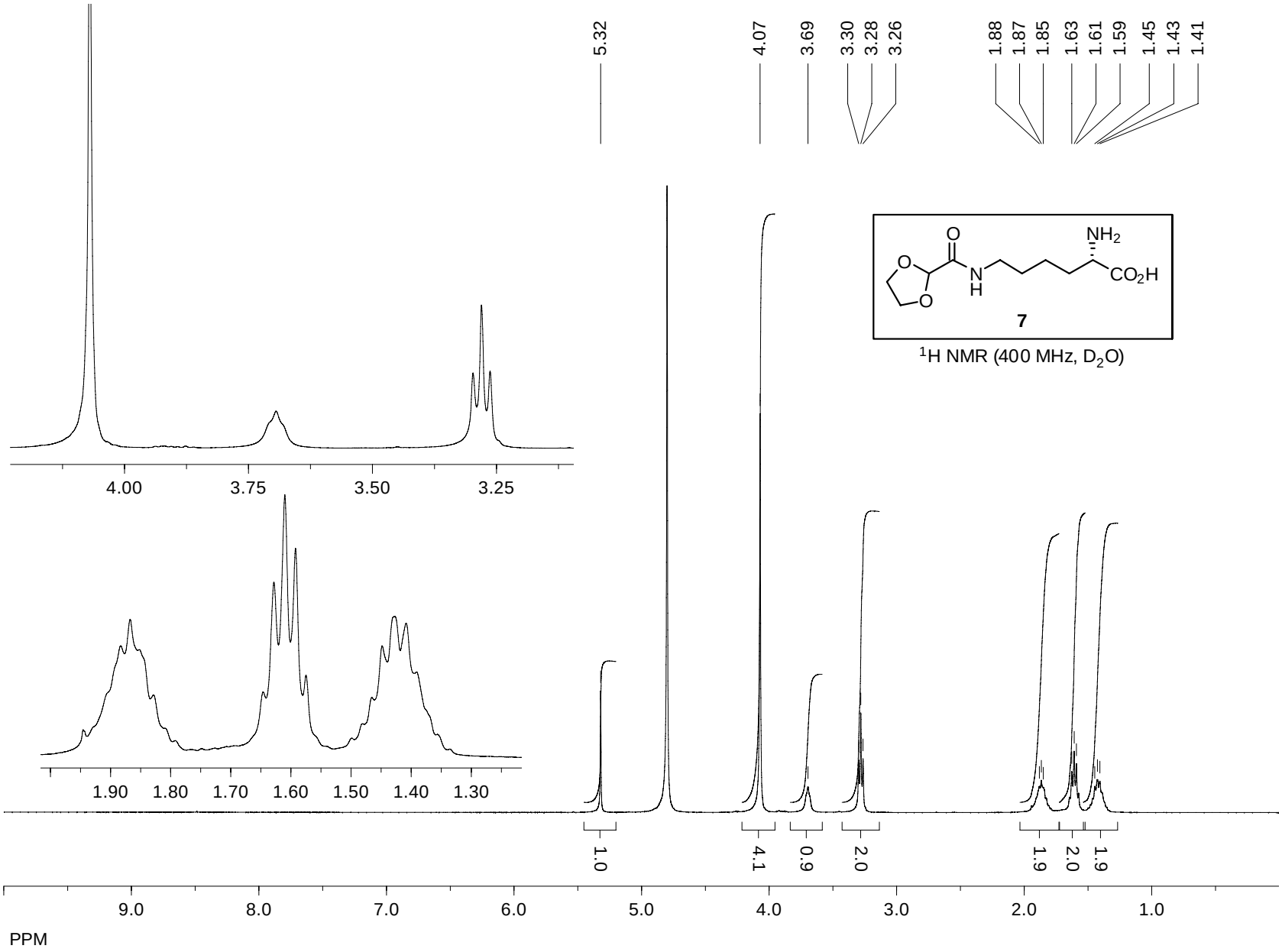
^{1}H-^{1}H COSY NMR (400 MHz, $CDCl_3$)

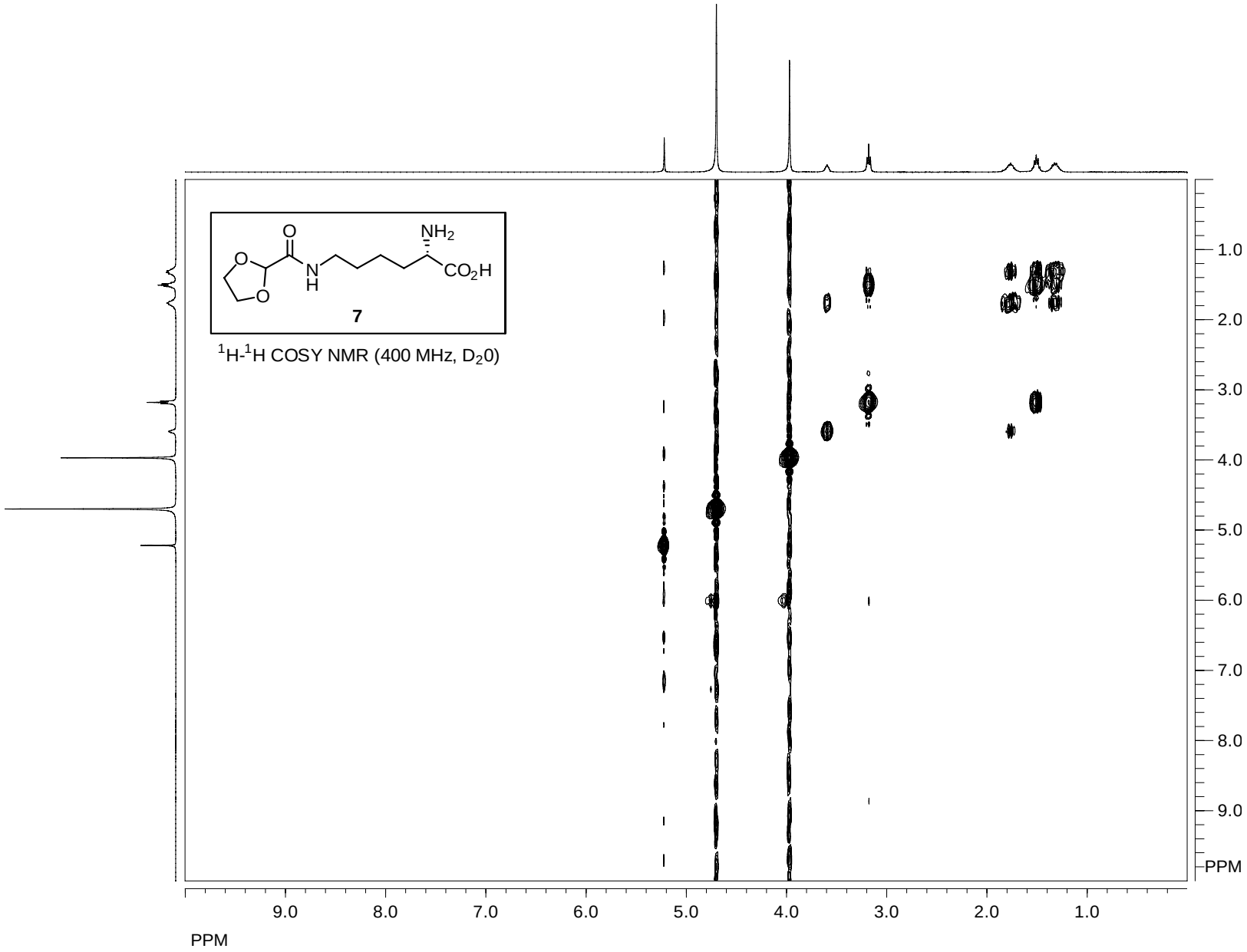


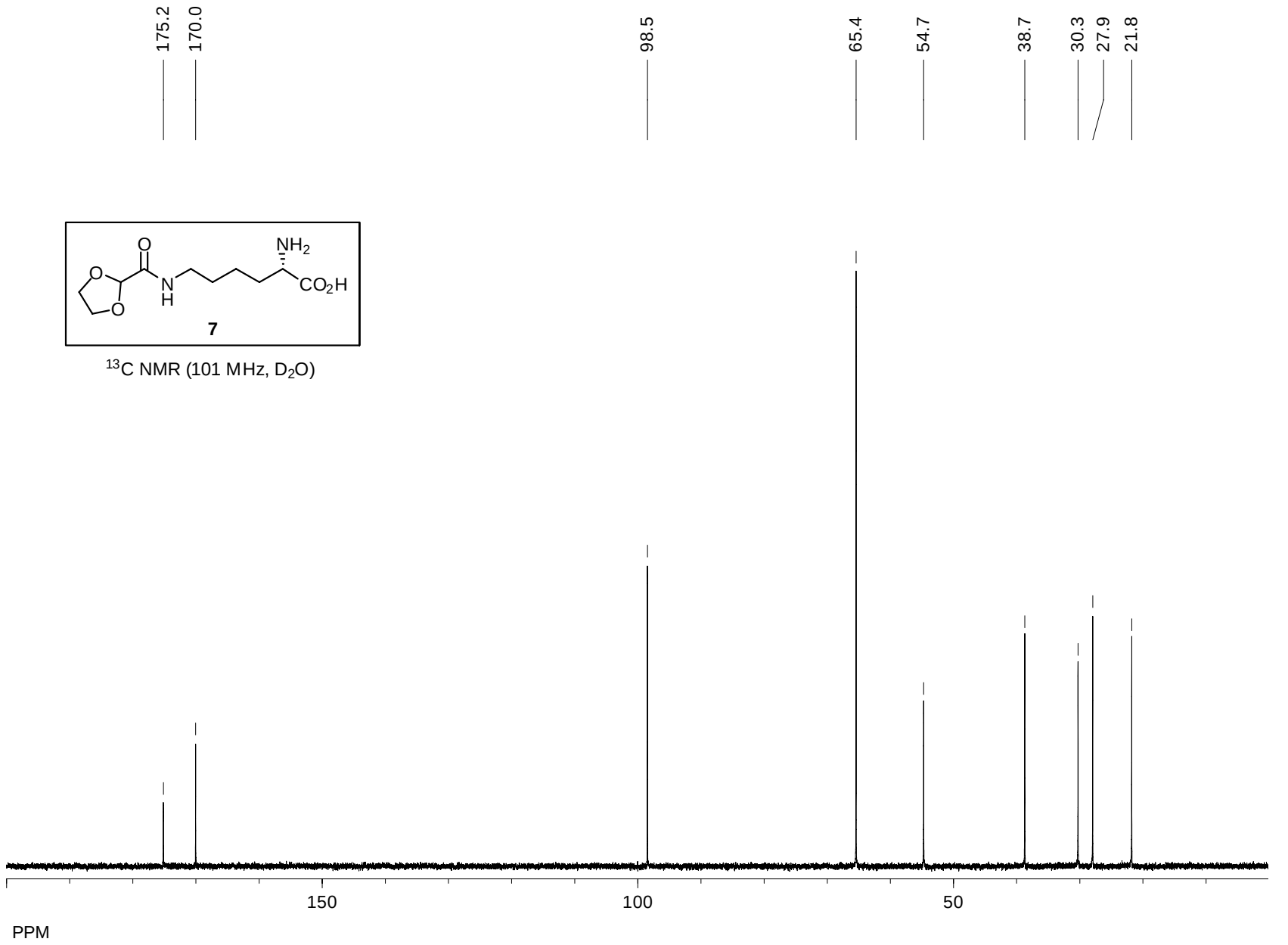
^{13}C NMR (101 MHz, $CDCl_3$)



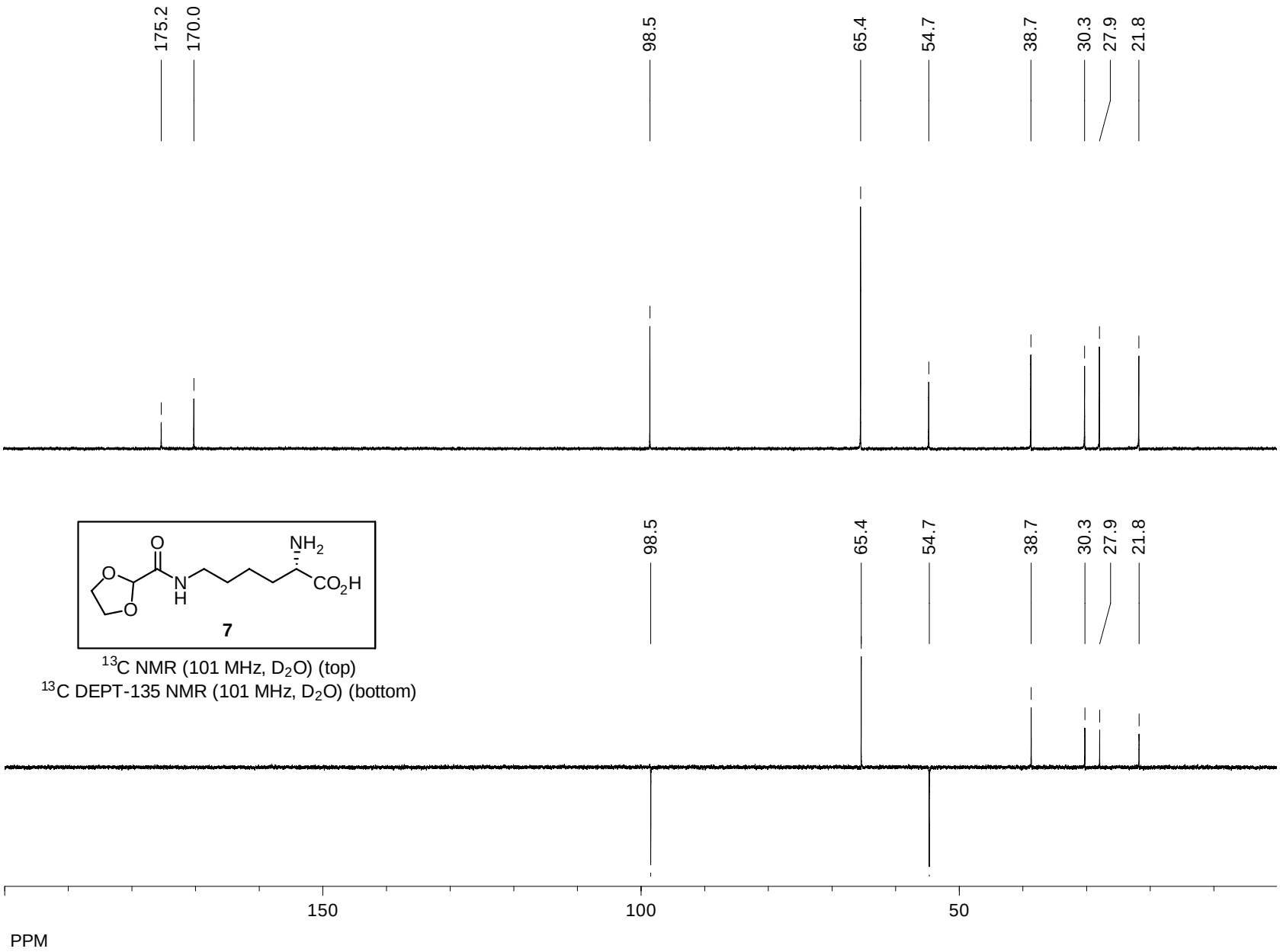
^{13}C NMR (101 MHz, $CDCl_3$) (top)
^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$) (bottom)


5.32
4.07
3.69
3.30
3.28
3.26
1.88
1.87
1.85
1.63
1.61
1.59
1.45
1.43
1.41
O
NH2
O
N
H
CO2H
O
7
1H NMR (400 MHz, D2O)
4.00
3.75
3.50
3.25
1.90
1.80
1.70
1.60
1.50
1.40
1.30
1.0
4.1
0.9
2.0
1.9
2.0
1.9
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

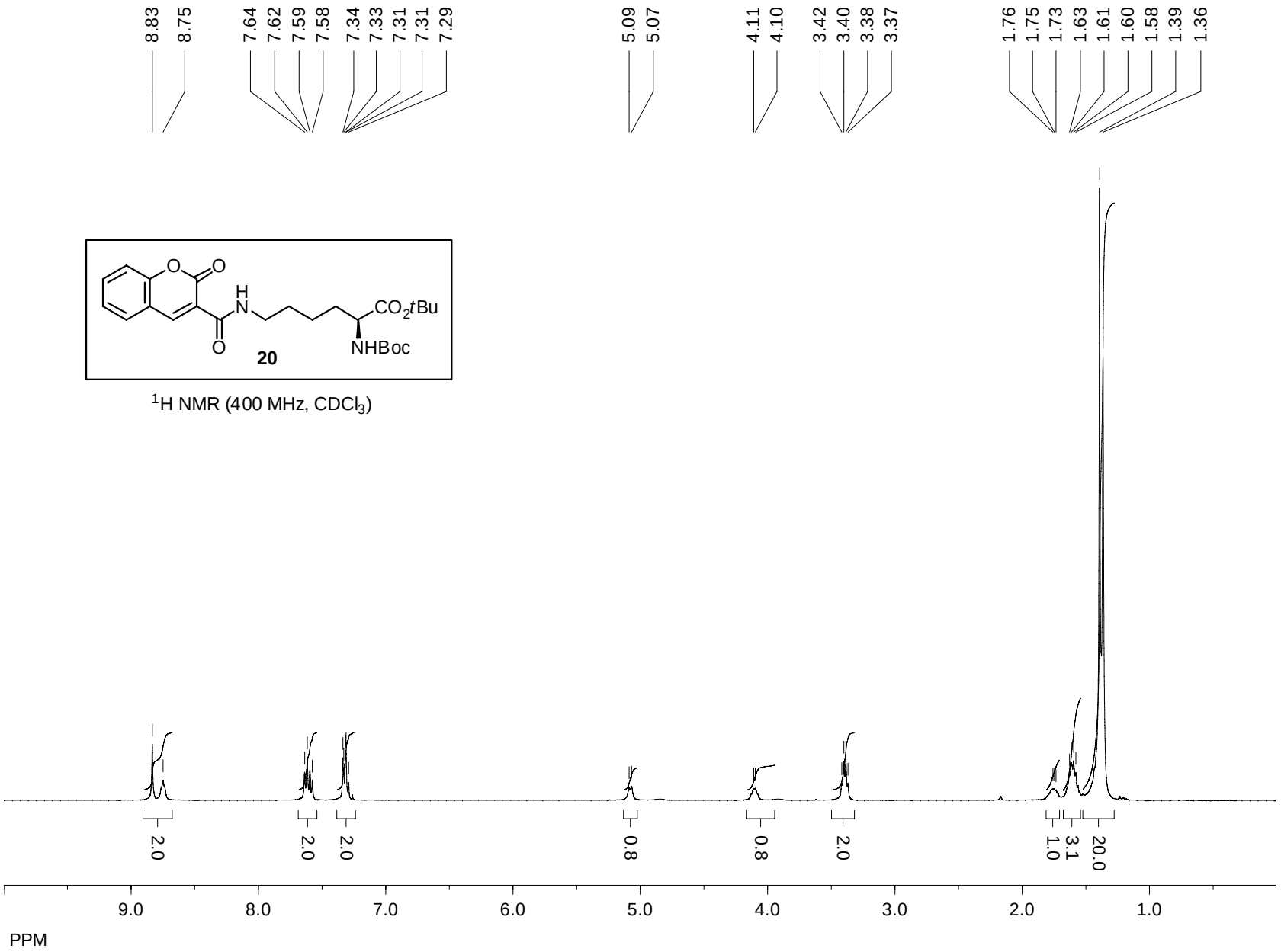

O
NH2
O
N
H
CO2H
O
7
1H-1H COSY NMR (400 MHz, D20)
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
PPM
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM



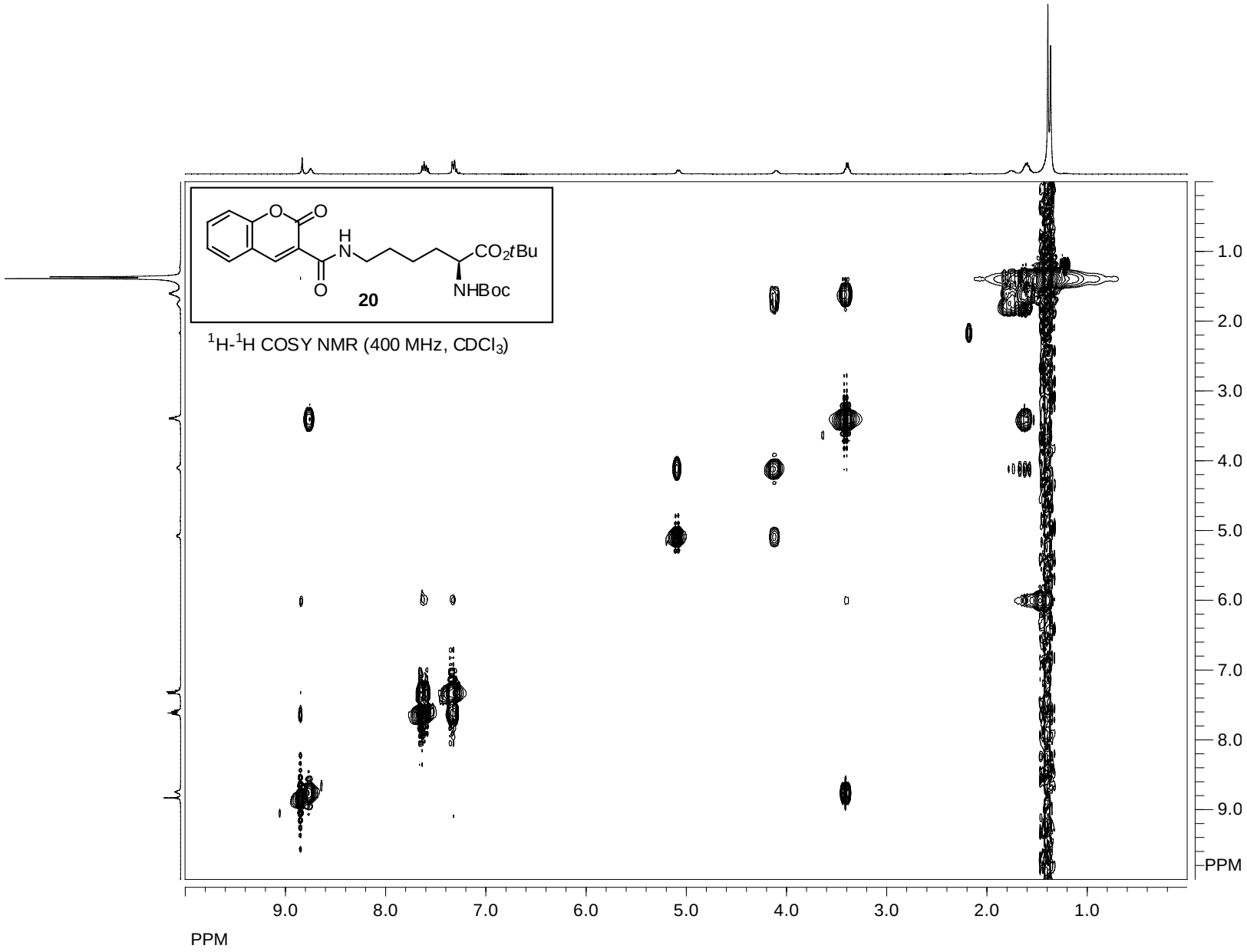
^{13}C NMR (101 MHz, D_2O)

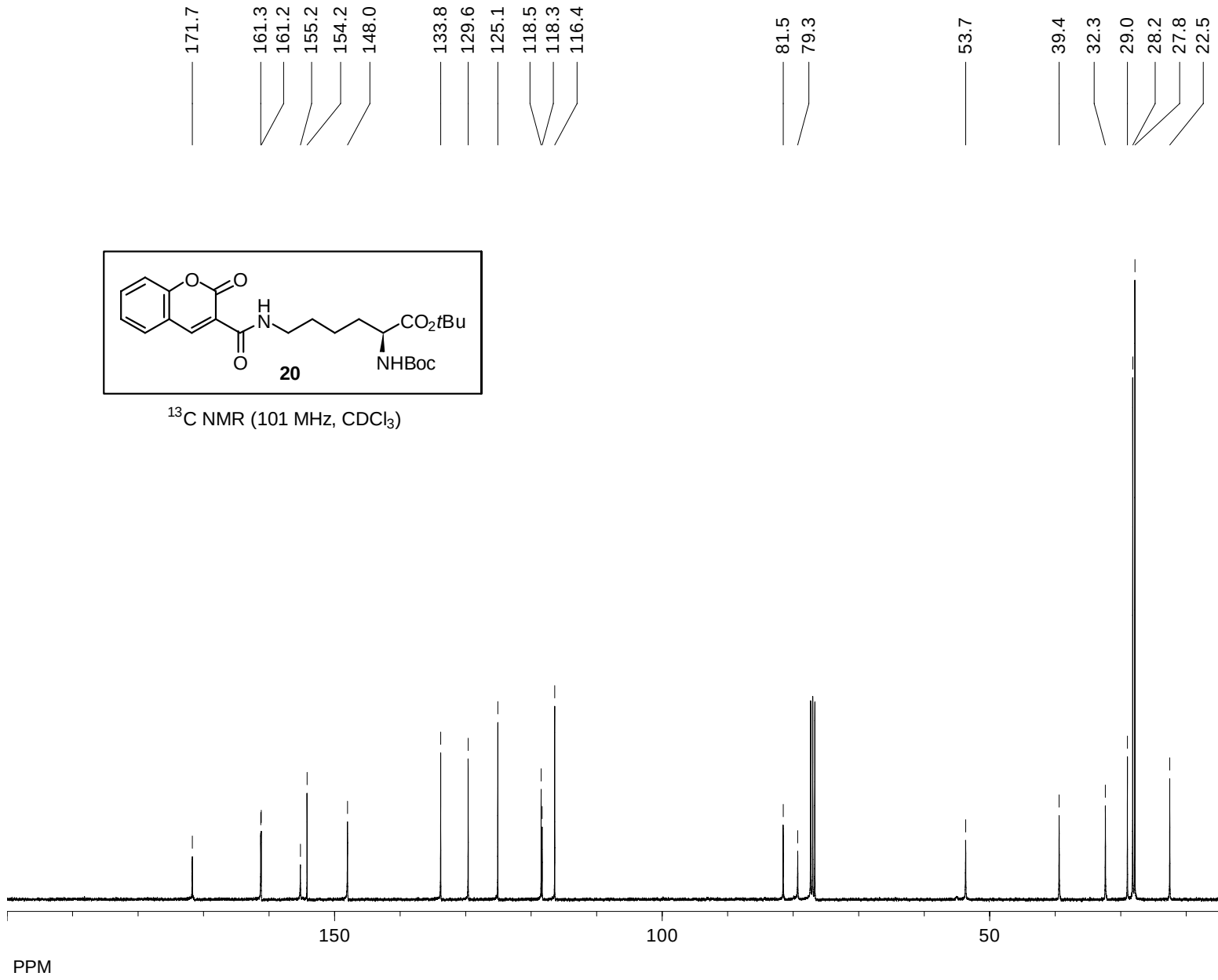


^{13}C NMR (101 MHz, D_2O) (top)
^{13}C DEPT-135 NMR (101 MHz, D_2O) (bottom)

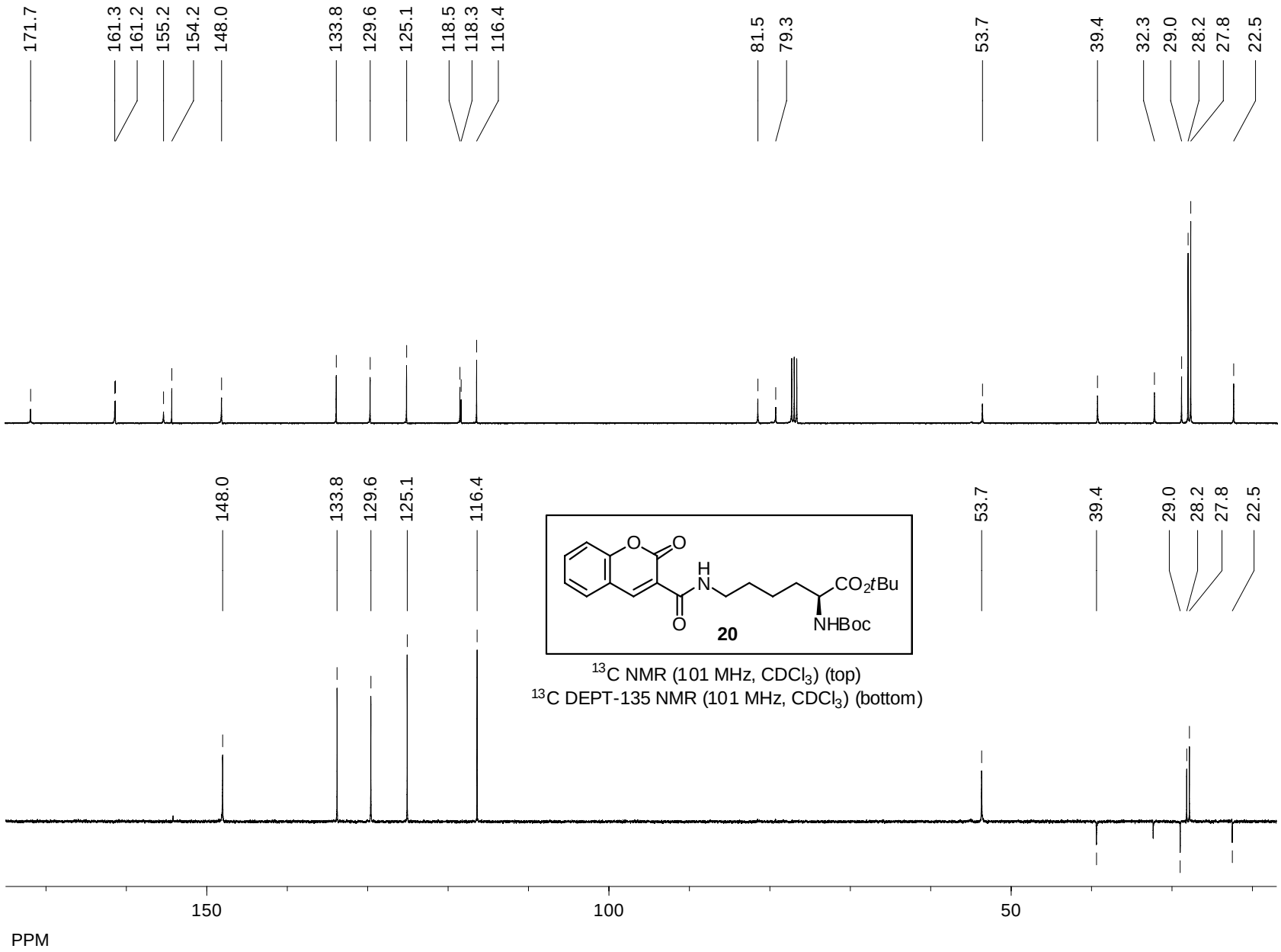


^{1}H NMR (400 MHz, $CDCl_3$)

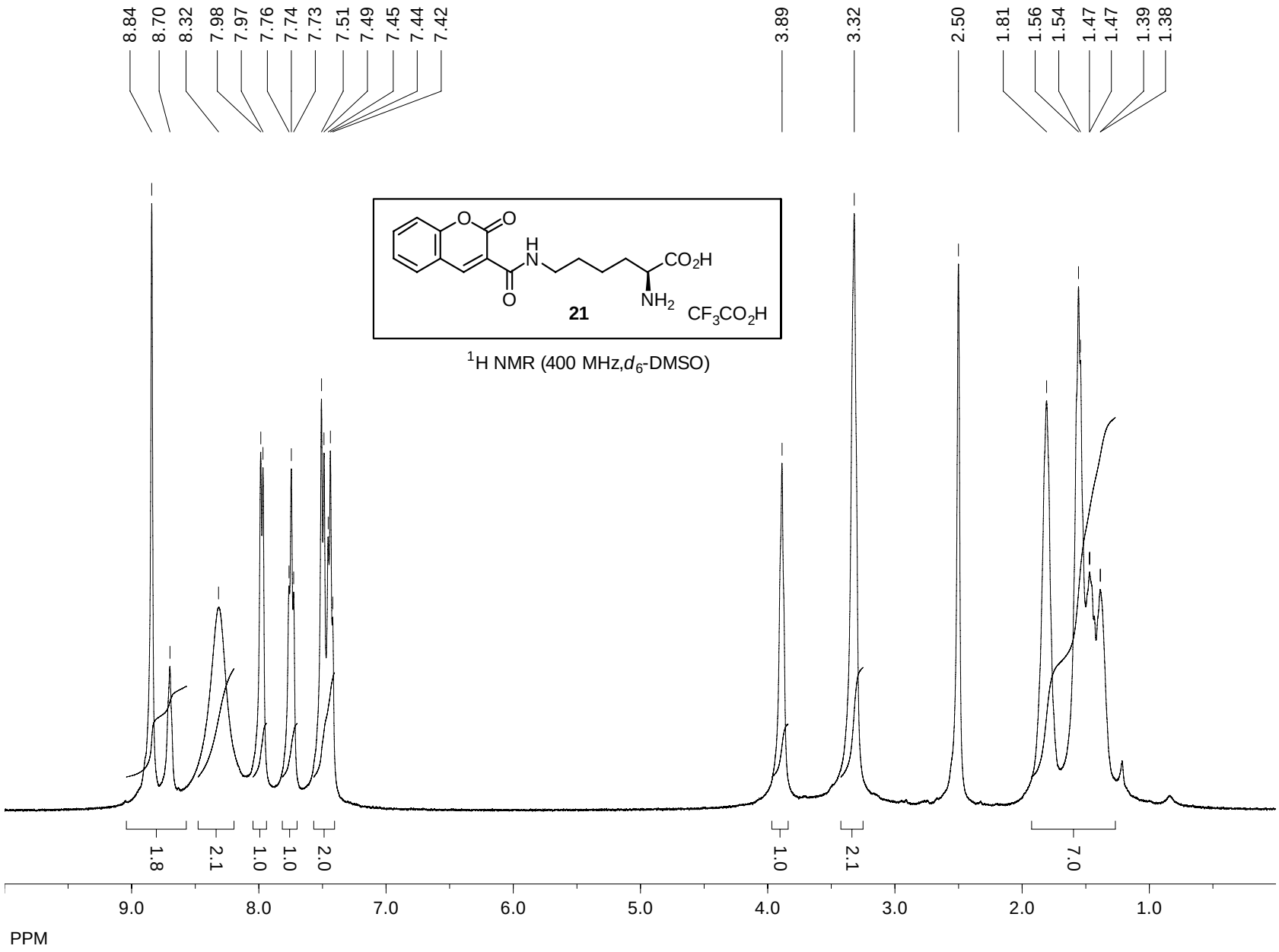

20
^{1}H-^{1}H COSY NMR (400 MHz, CDCl$_3$)
PPM
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0

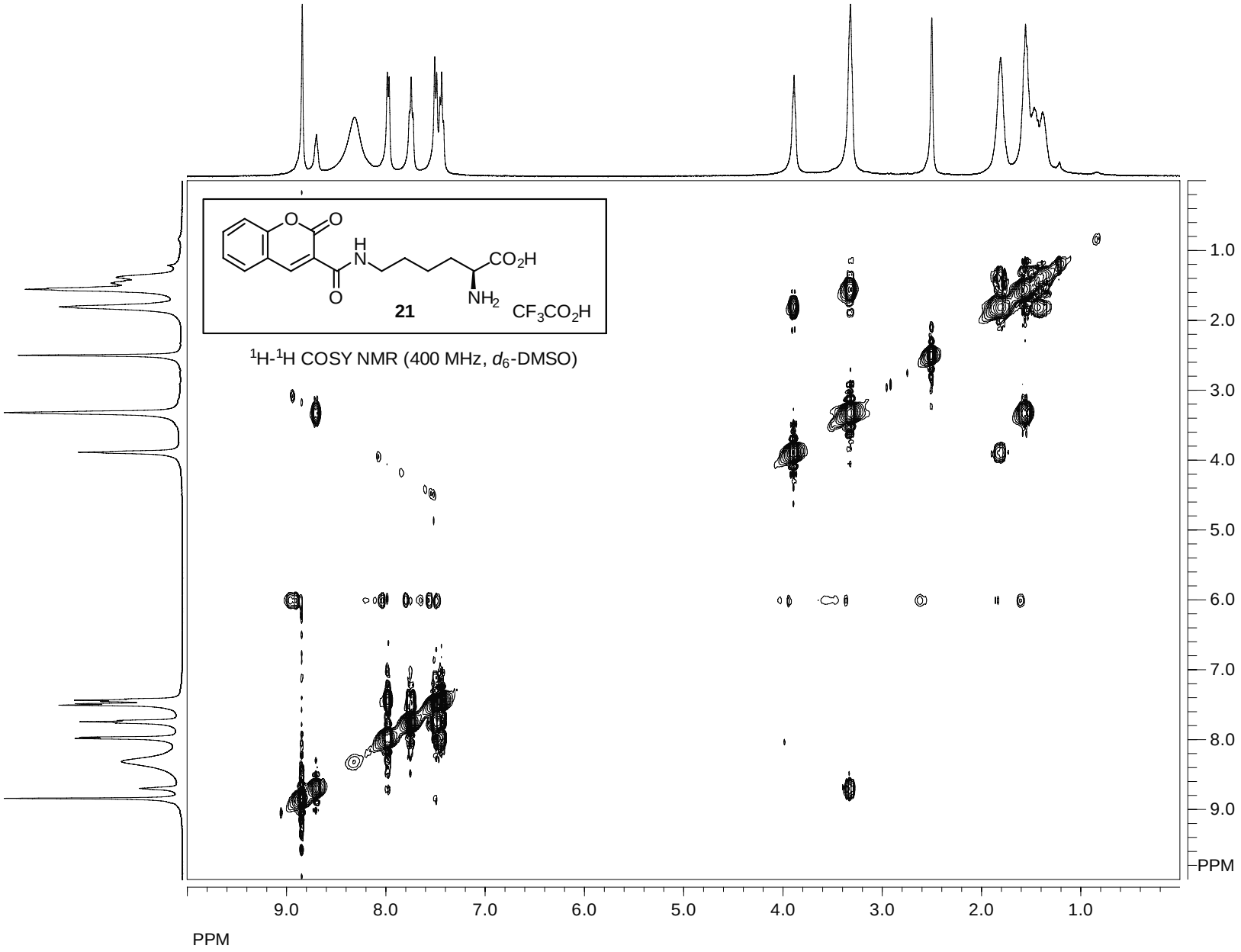


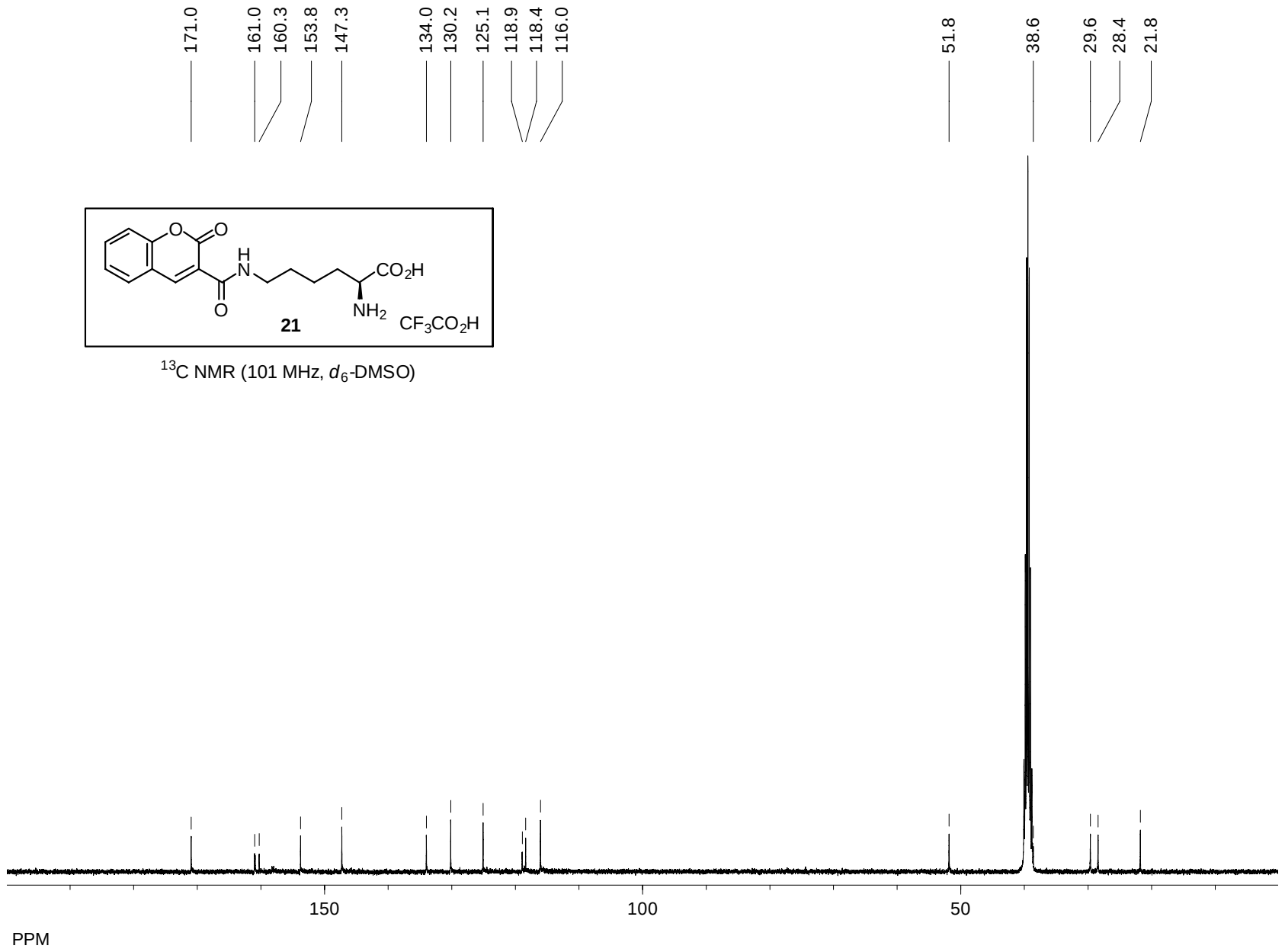
^{13}C NMR (101 MHz, $CDCl_3$)



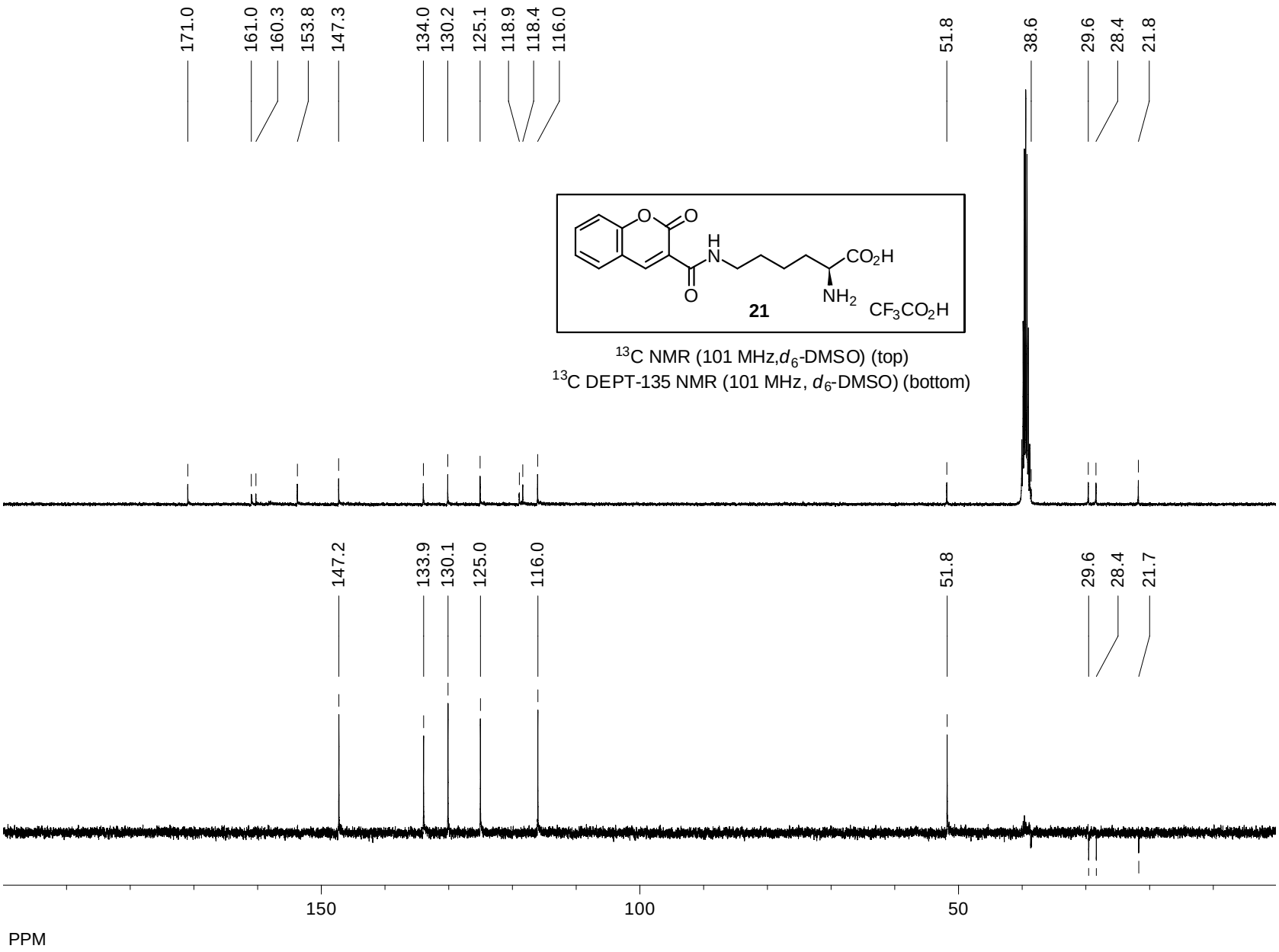
^{13}C NMR (101 MHz, $CDCl_3$) (top)
^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$) (bottom)


8.84
8.70
8.32
7.98
7.97
7.76
7.74
7.73
7.51
7.49
7.45
7.44
7.42
3.89
3.32
2.50
1.81
1.56
1.54
1.47
1.47
1.39
1.38
21
CF3CO2H
1H NMR (400 MHz,d6-DMSO)
1.8
2.1
1.0
1.0
2.0
1.0
2.1
7.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

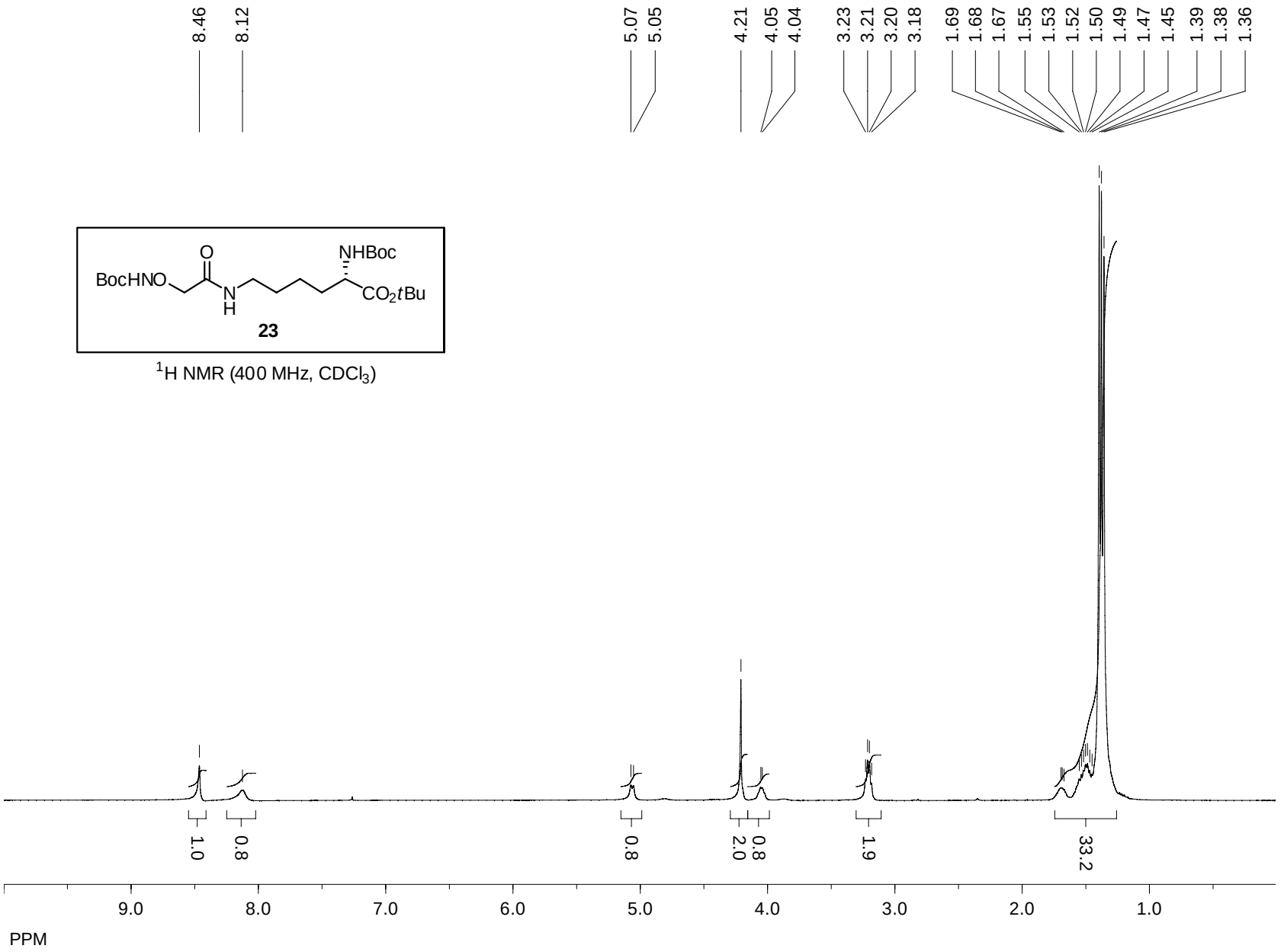

21
CF3CO2H
1H-1H COSY NMR (400 MHz, d6-DMSO)
PPM
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0



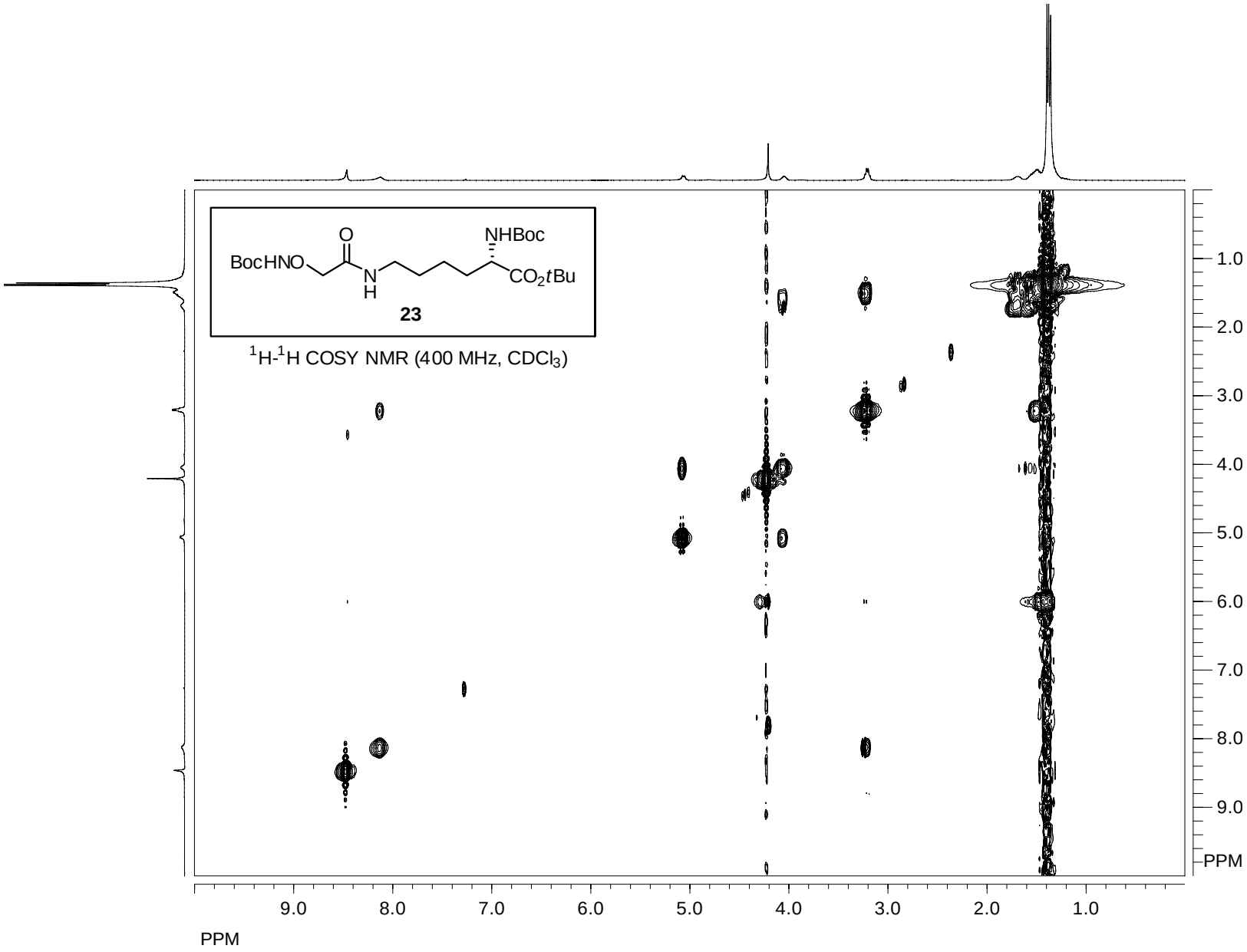
^{13}C NMR (101 MHz, d_6-DMSO)

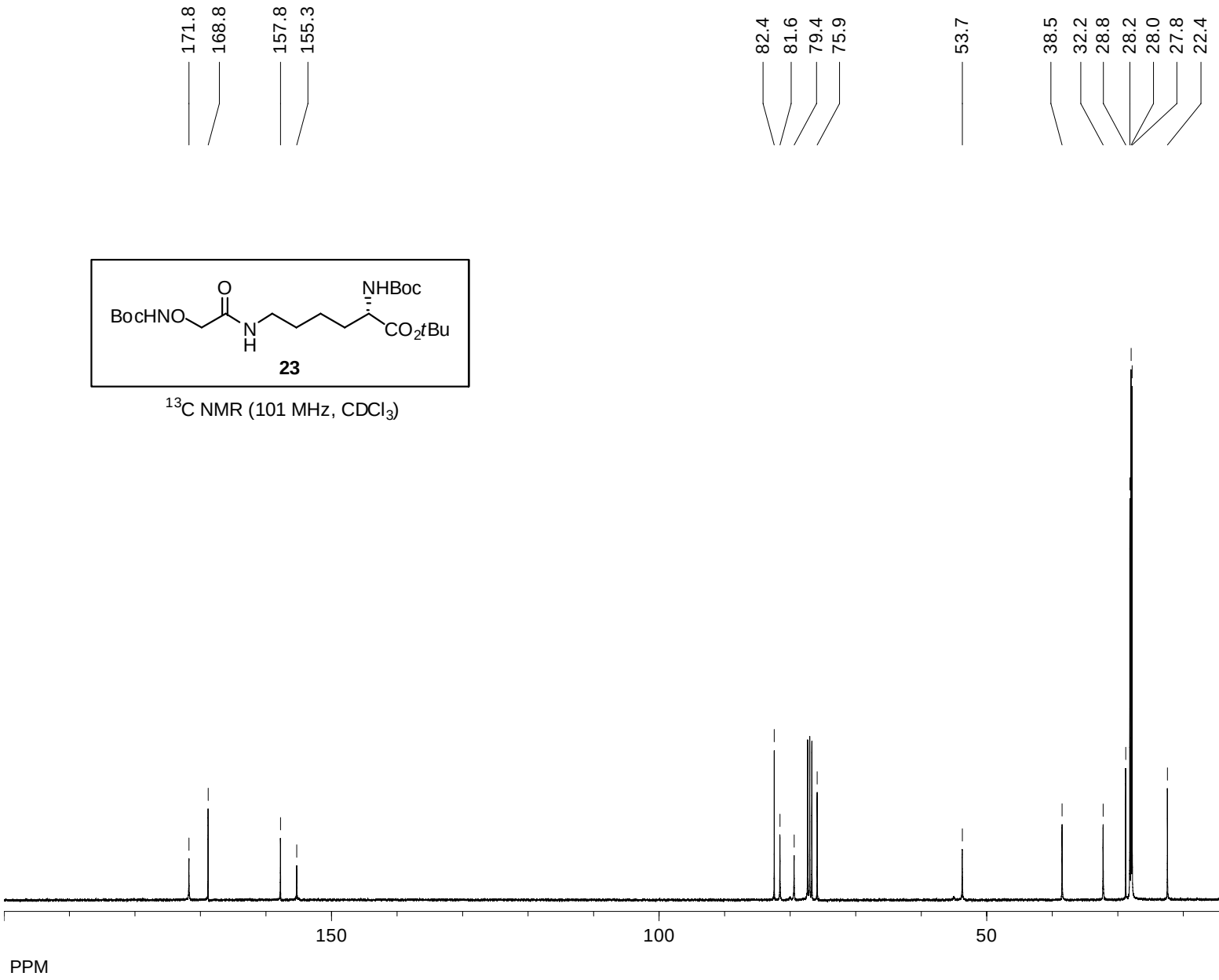


^{13}C NMR (101 MHz,d_6-DMSO) (top)
^{13}C DEPT-135 NMR (101 MHz, d_6-DMSO) (bottom)



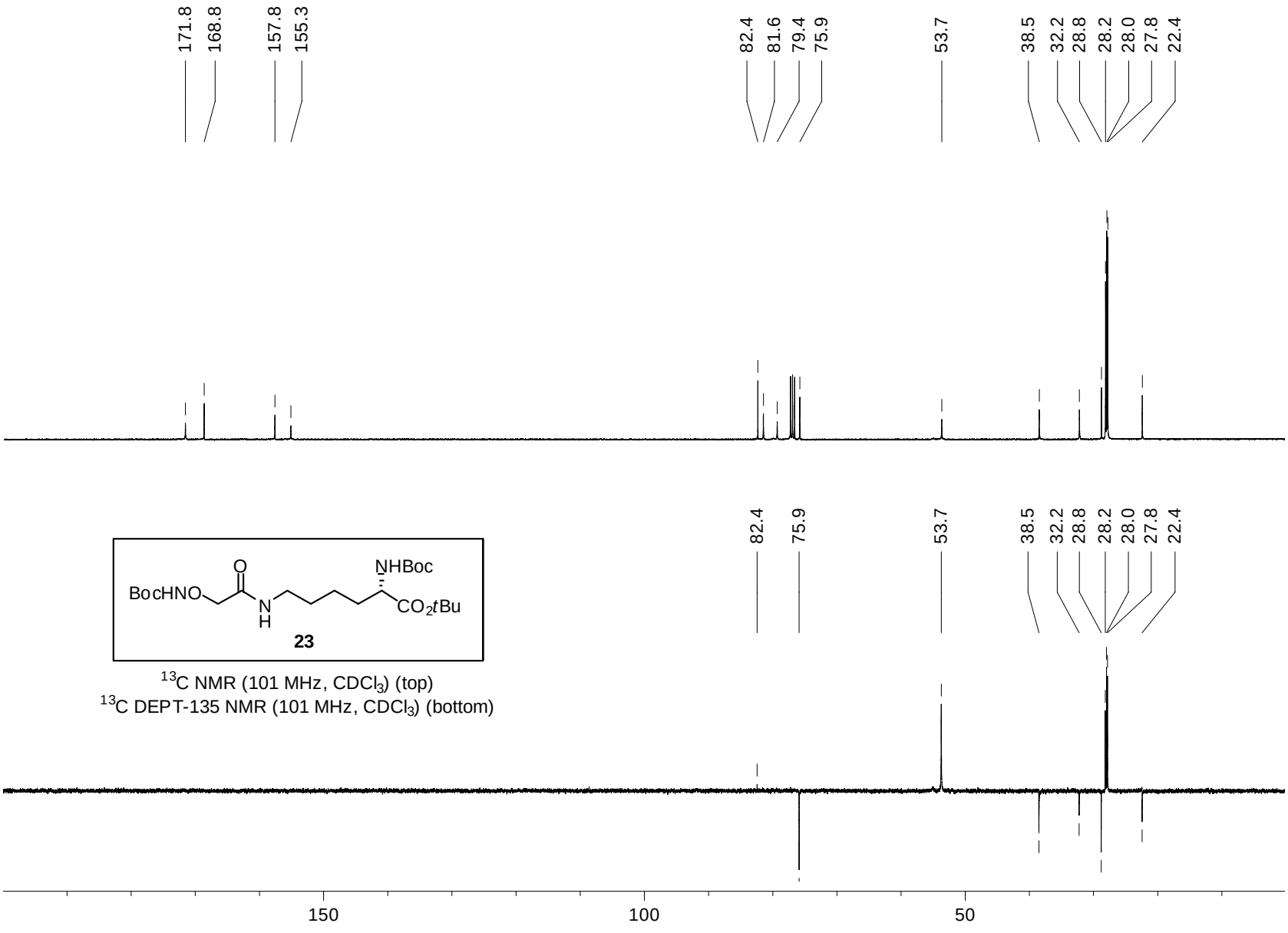
^{1}H NMR (400 MHz, $CDCl_3$)


BocHNO
O
N
H
NHBoc
CO2tBu
23
1H-1H COSY NMR (400 MHz, CDCl3)
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM

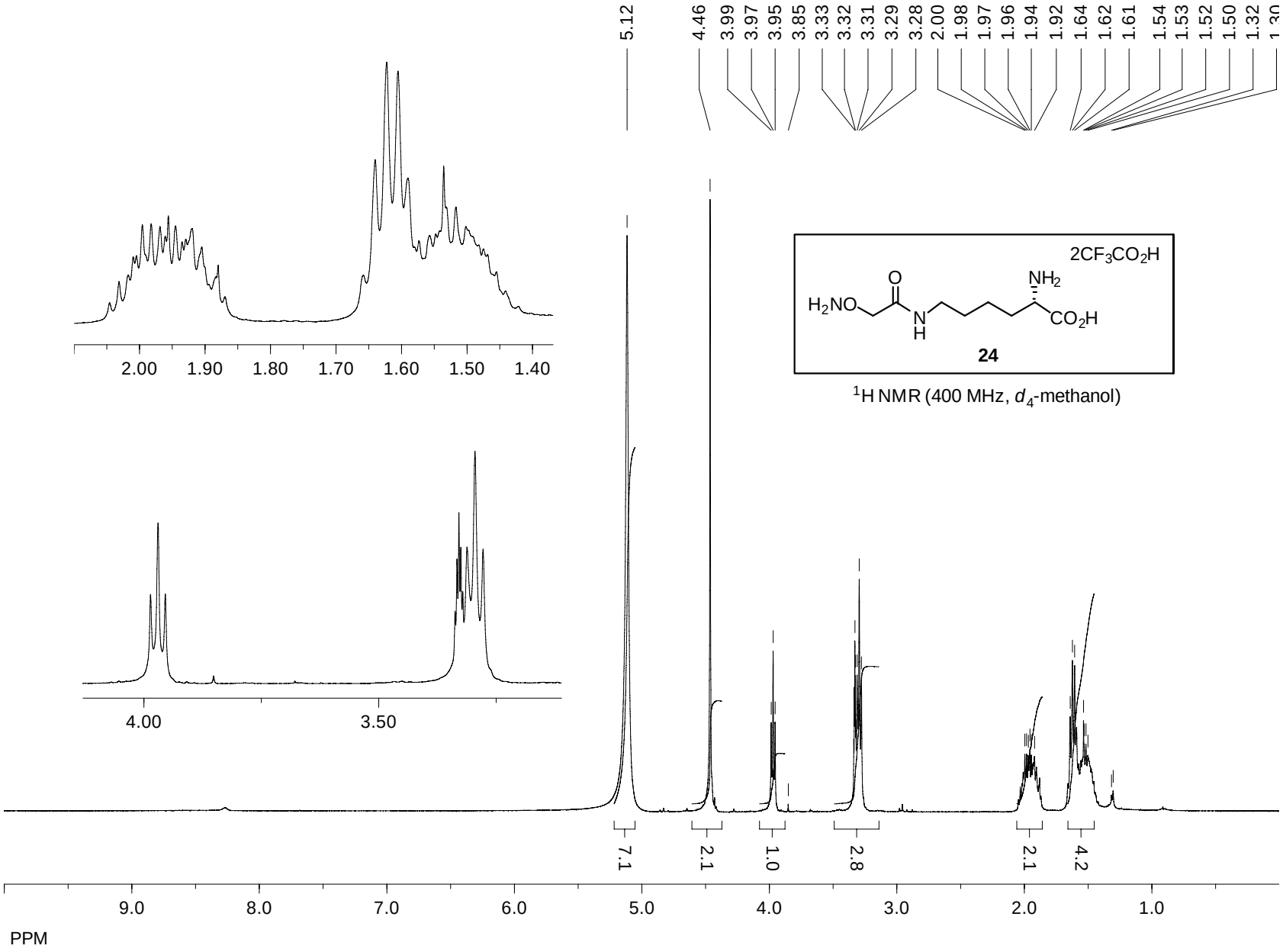


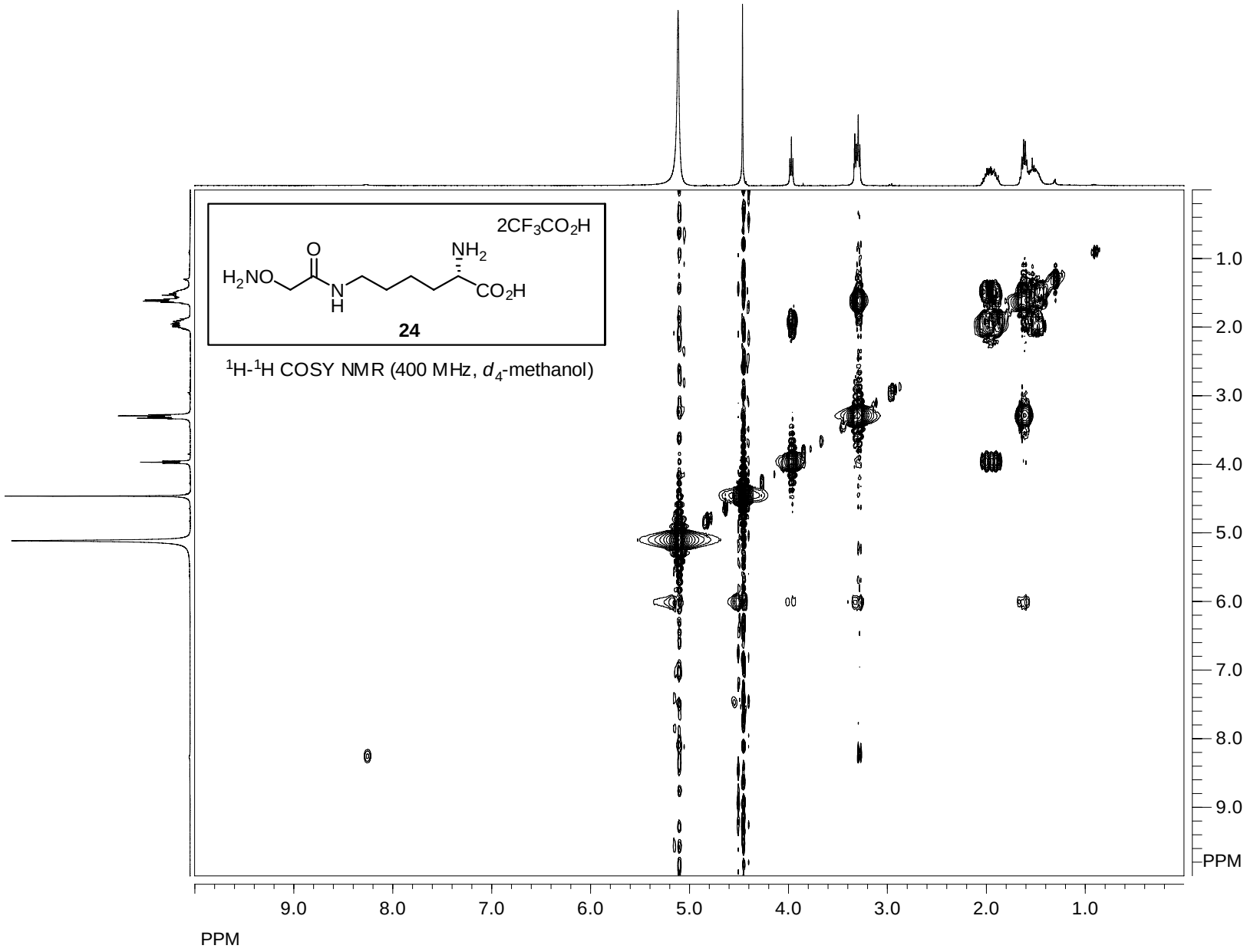
23

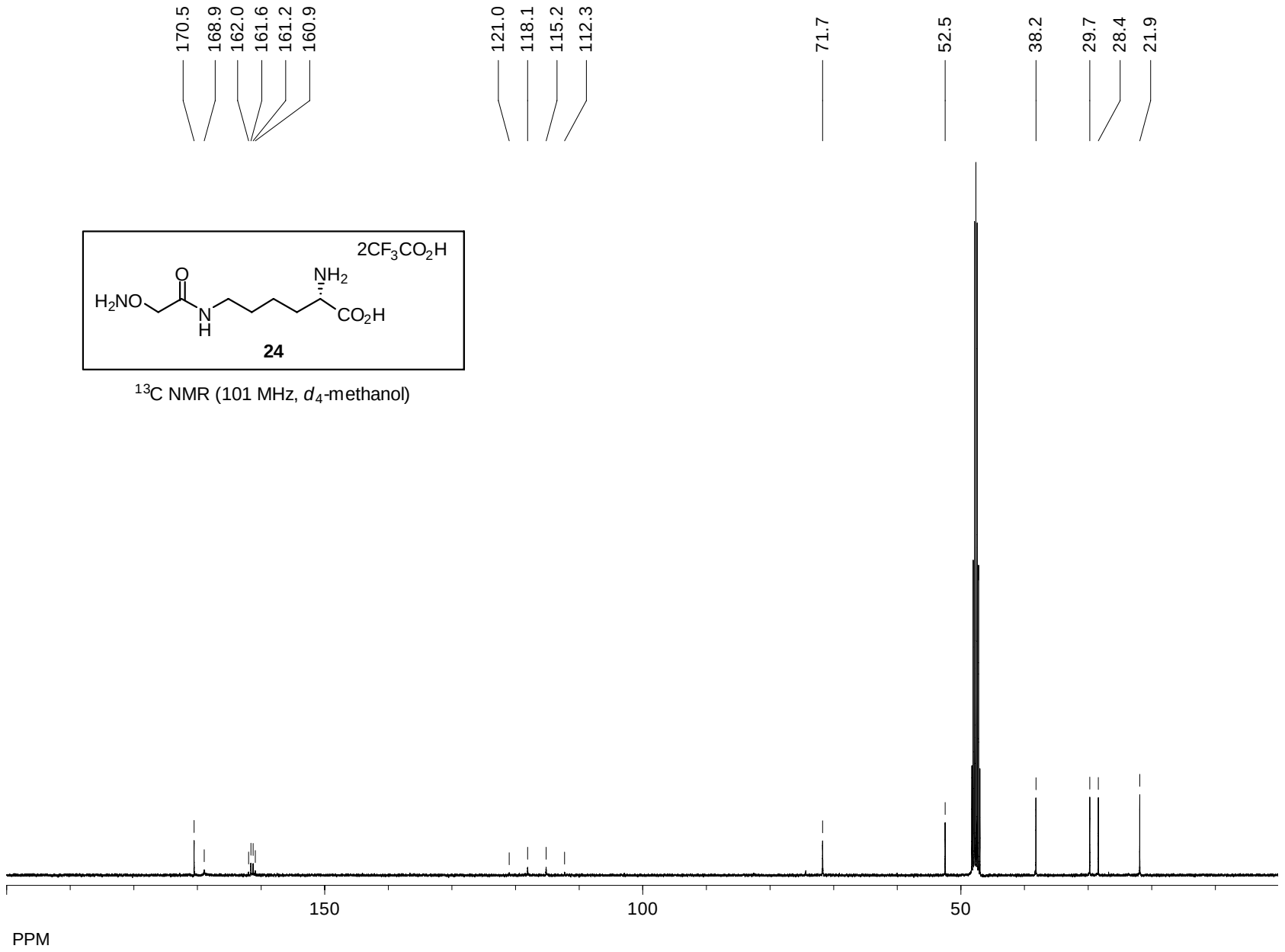
^{13}C NMR (101 MHz, $CDCl_3$)



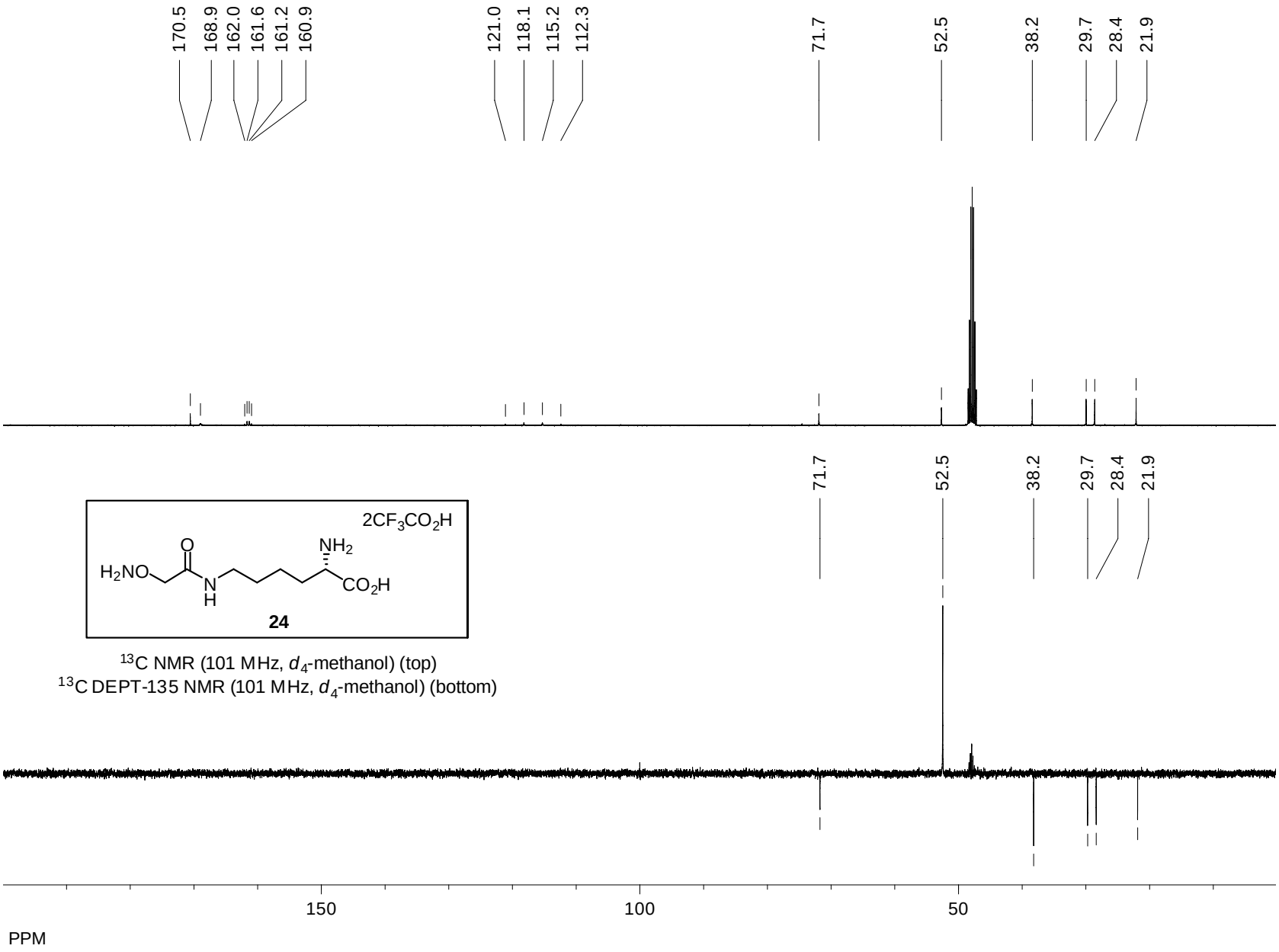
^{13}C NMR (101 MHz, $CDCl_3$) (top)
^{13}C DEPT-135 NMR (101 MHz, $CDCl_3$) (bottom)


5.12
4.46
3.99
3.97
3.95
3.85
3.33
3.32
3.31
3.29
3.28
2.00
1.98
1.97
1.96
1.94
1.92
1.64
1.62
1.61
1.54
1.53
1.52
1.50
1.32
2.00 1.90 1.80 1.70 1.60 1.50 1.40
4.00 3.50
2CF3CO2H
H2NO
O
N
H
NH2
CO2H
24
1H NMR (400 MHz, d4-methanol)
7.1
2.1
1.0
2.8
2.1
4.2
9.0 8.0 7.0 6.0 5.0 4.0 3.0 2.0 1.0
PPM


2CF3CO2H
O
NH2
H2NO
N
H
CO2H
24
1H-1H COSY NMR (400 MHz, d4-methanol)
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
PPM
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
PPM



^{13}C NMR (101 MHz, d_4-methanol)



^{13}C NMR (101 MHz, d_4-methanol) (top)
^{13}C DEPT-135 NMR (101 MHz, d_4-methanol) (bottom)